

ECOLOGY AND POPULATION STRUCTURE OF *VIBRIONACEAE* IN THE COASTAL
OCEAN

By
Sarah Pacocha Preheim
B.S., Carnegie Mellon University, 1997

Submitted to the Department of Civil and Environmental Engineering, Massachusetts Institute of
Technology and the Woods Hole Oceanographic Institution in Partial Fulfillment of the
Requirements for the Degree of

Doctor of Philosophy in the Field of Biological Oceanography

at the
MASSACHUSETTS INSTITUTE OF TECHNOLOGY
and the
WOODS HOLE OCEANOGRAPHIC INSTITUTION
February 2010

© 2010 Massachusetts Institute of Technology and Woods Hole Oceanographic Institution
All rights reserved

Signature of Author

Joint Program in Oceanography
Department of Civil and Environmental Engineering
Massachusetts Institute of Technology
and Woods Hole Oceanographic Institution
January 12th, 2010

Certified by

Martin F. Polz
Professor of Civil and Environmental Engineering
Massachusetts Institute of Technology
Thesis Supervisor

Accepted by

Simon Thorrold
Chair, Joint Committee for Biological Oceanography
Woods Hole Oceanographic Institution

Accepted by

Daniele Veneziano
Chairman, Departmental Committee for Graduate Students
Massachusetts Institute of Technology

ECOLOGY AND POPULATION STRUCTURE OF *VIBRIONACEAE* IN THE COASTAL OCEAN

by

Sarah P. Preheim

Submitted to the Department of Civil and Environmental Engineering, Massachusetts Institute of Technology and the Woods Hole Oceanographic Institution on Jan. 12th, 2010
in Partial Fulfillment of the Requirements for the Degree of Doctor of Philosophy in the
Field of Biological Oceanography

Abstract

Extensive genetic diversity has been discovered in the microbial world, yet mechanisms that shape and maintain this diversity remain poorly understood. This thesis investigates to what extent populations of the gamma-proteobacterial family, *Vibrionaceae*, are ecologically specialized by investigating the distribution across a wide range of environmental categories, such as marine invertebrates or particles in the water column. Additionally, it seeks to determine whether *in situ* population distributions directly result from a competitive advantage over other *Vibrio* populations. This was investigated by *in vitro* competition assays on mixtures of native, sterilized particles. Generalist populations were found to dominate the associations with marine invertebrates, consistent with a model of high migration dominated population assembly. A majority of populations occurred broadly within and among the different types of invertebrates sampled, with one population being a near perfect generalist with regard to seasons, host taxa and body regions. High variability across host individuals, consistent with a scenario of stochastic clonal expansion, was especially pronounced in crab and zooplankton samples. Specialization, demonstrated by specific and reproducible association with different particle types in the water column, is more common than specialization within invertebrate hosts. Co-existing *Vibrio* species show strong preferences for different types of particulate matter in the water column suggesting that competition for limited resources influences their evolution. While populations show different growth profiles on particle derived substrates, relative growth advantages of specialist populations in competition with other *Vibrio* populations on native particles may not be sufficient to explain observed environmental distributions. Instead, populations may gain an advantage on these particles by colonizing the living plant or zooplankton prior to death and degradation into particulate matter. In summary, although vibrios are known commensals of marine invertebrates, evidence suggests that population structure within animals is fairly weak compared to suspended particles in the water column. This highlights the importance of comparing multiple environmental categories and migration among them to investigate population structure and adaptation.

Thesis Supervisor: Martin F. Polz

Title: Professor of Civil and Environmental Engineering

ACKNOWLEDGMENTS:

Special thanks go to Martin Polz for being a wonderful mentor, providing insight, support, understanding and humor. His countless comments on my chapters and proposals have helped me better express and focus my scientific ideas. He granted me the creative freedom to design and perform experiments, while providing sufficient guidance to keep my research focused and informative. Polz lab is filled with interesting and intelligent people, and Martin deserves credit for cultivating a creative and collaborative lab environment. I'm very grateful to Eric Alm for all of his help. Eric has been an excellent sounding board for new ideas. I'm also glad to have received supportive and insightful comments from Ed DeLong, Dan Weinreich and Stefan Sievert made during committee meeting and other correspondences.

This work would not be possible without the help and support of current and former Polz lab mates. Many of them have, in one way or another, made experimentation easier by helping me collect samples and analyze data and building a knowledge base for me to work with. Former members, Janelle Thompson, Dana Hunt and Vanja Klepac-Ceraj provided a great deal of support when I first arrived in the lab, encouraging me and making me feel welcome. Current members of the Polz lab have been great when I needed help or encouragement as well. Although processing water samples for 24 hrs in a row, or filtering water until late into the night can be very trying, everyone tackles these tasks with more cheer than can reasonably be expected. Yan Boucher and Hans Wildschutte were especially helpful while collecting invertebrate and particle samples. I could not have sampled the environment for vibrios so extensively without their help. Alison Takemura was also a tremendous help and became very adept with a hammer while extracting DNA samples for the *in vitro* competition assays. I'm very glad that she was willing to lend a helping hand with those assays. Almost every member of the Polz lab helped me at one point or another, whether sampling or streaking and characterizing strains. I was also able to work with three amazing undergraduate students as part of the UROP program; Kelsey Byers, Shelley Gu and Grace Yuen. Although their work did not directly contribute to this thesis, their results created the foundation for this thesis. They all showed exceptional potential as researchers and scientists. Additionally, it has been a tremendous privilege to be able to walk down the hall to get advice, reagents or equipment from the Chisholm, DeLong, Alm and Thompson labs and lab members, which has made my time at MIT much more enjoyable.

My husband Chris and my daughter Natalie have been a crucial part of my success, providing me with love, support and much needed diversions from my lab activities. My sisters and parents have also shown so much love and support for me and I'm glad for their constant encouragement and sagely advice. Without the love and support of my entire family, I would not have had the opportunity to do what I love and pursue a career that I was drawn to.

I would like to thank the Linden Fellowship for their support during the Fall, 2003- Spring, 2004 and the Martin Family Society of Fellows for Sustainability for their support during the Fall 2005. Work in this thesis was also made possible by support from the following agencies: Woods Hole Center for Oceans and Human Health, which is

funded jointly by the National Science Foundation (OCE-0430724) and the National Institute for Environmental Health Sciences (P50 ES012742); DOE-Chicago (DE-FG02-08ER64516); The Gordon and Betty Moore Foundation (Letter Agreement 1108).

TABLE OF CONTENTS:

Abstract	3
Acknowledgments	5
List of Figures	8
List of Tables	11
Chapter 1: Introduction	13
Chapter 2: Generalists Dominate Vibrio Populations in the Microbiome of Marine Animals	35
Chapter 3: Resource Partitioning and Substrate Pre-Colonization Likely Influence Vibrio Population Structure on Water Column Particles	67
Chapter 4: Reproducibility and Specificity of Vibrio Population Associations with Environmental Parameters	121
Chapter 5: Conclusions and Future Work	173

LIST OF FIGURES:

Chapter 1

- Fig. 1 Model of bacterial evolution resulting in low effective population sizes (bottleneck), where concentrated patches of nutrients are important in creating population structure. 33

Chapter 2

- Fig. 1 Sampling strategy for comparing population structure of *Vibrionaceae* bacteria across animal hosts and water column samples in two seasons (fall and spring). 59

- Fig. 2 Population structure predictions of *Vibrionaceae* bacteria recovered from different types of animal samples. 60

- Fig. 3 Diversity of population fractions among different host specimens. 62

- Fig. S1 Comparison of phylogenetic relationships of *V. splendidus*-like populations #18-25 from Hunt, David et al. 2008 and population #16 from this study. 64

- Fig. S2. Comparison of number of strains per sequence type (ST) for each sample of crab and mussel compartments, aggregate zooplankton, and seawater. 65

Chapter 3

- Fig. 1. Predicted populations, habitat preferences and environmental distributions for strains isolated from environmental sampling. 103

- Fig. 2. *Vibrio* populations associated with brown algae (*Fucus vesiculosus*), one of the likely sources of plant-derived particles. 104

- Fig. 3. Relative abundance of two strains with different predicted particle preferences on plant-derived (a) and zooplankton particles (b) or seawater alone (c). 105

- Fig. 4. Relative abundance of two populations (*Vibrio* sp. F10 and *V. breoganii*) with different predicted particle preferences on plant-derived (a, d, g) or zooplankton particles (b, e, h) or seawater alone (c, f, i). 106

Fig. 5. Outcomes for individual strain-pair competition experiments with <i>Vibrio</i> sp. F10 (zooplankton specialist) and <i>V. breoganii</i> (plant-derived particle specialist) on zooplankton particles.	108
Fig. 6. Relative abundance of two populations (<i>Vibrio</i> sp. F10 and <i>V. crassostreae</i>) with different predicted particle preferences on plant-derived (a,d,g) or zooplankton particles (b,e,h) or seawater alone (c,f,i).	110
Fig. 7. Strain to strain variation in outcome in competition experiments with <i>Vibrio</i> sp. F10 (zooplankton specialist) and <i>V. crassostreae</i> (plant-derived particle specialist) on zooplankton particles.	112
Fig. 8. Relative abundance of two populations (<i>V. breoganii</i> and <i>V. crassostreae</i>) with the same predicted particle preferences on plant-derived (a,d,g) or zooplankton particles (b,e,h) or seawater alone (c,f,i).	114
Chapter 4	
Fig. 1. Phylogenetic relationship of populations predicted across three studies of <i>Vibrio</i> population structure (fract-2006, invert-2007 and particle-2007, explained in the text) inferred by maximum likelihood using <i>hsp60</i> sequences.	147
Fig. 2. Maximum likelihood estimation of phylogenetic relationship between named species and population representatives using partial <i>atpA</i> sequences.	149
Fig. 3. Maximum likelihood estimation of phylogenetic relationship between named species and population representatives using partial <i>gyrB</i> sequences.	151
Fig. 4. Maximum likelihood estimation of phylogenetic relationship between named species and population representatives using partial <i>pyrH</i> sequences.	153
Fig. 5. Maximum likelihood estimation of phylogenetic relationship between named species and population representatives using partial <i>recA</i> sequences.	155
Fig. 6. Maximum likelihood estimation of phylogenetic Relationship between named species and population representatives using partial <i>topA</i> sequences.	157

Fig. 7. Agglomerative hierarchical process used to generate a dendrogram (left) ordering populations by environmental similarity.

159

LIST OF TABLES:

Chapter 2

Table S1. Reproducibility of the total number of habitats resulting from one hundred independent trials.	66
--	----

Table S2. Reproducibility of inferred habitats.	66
---	----

Chapter 3

Table 1. Differential growth on various carbon sources between particle type specialists.	116
---	-----

Table 2. Distribution of plant-particle specialist populations across eight replicate samples of plant-derived particles.	117
---	-----

Table 3. Antagonistic interactions between strains used in <i>in vitro</i> competition assays, including both growth characteristics and directed inhibition.	117
---	-----

Table 4. Temperature and Salinity Conditions During Spring and Fall Sampling Dates	118
--	-----

Table 5. Distribution of habitat predictions over 100 replicate iterations of the clustering algorithm (AdaptML) representing relative habitat stability.	118
---	-----

Table 6. Strain combinations for <i>in vitro</i> competition assays including initial inoculation sources	119
---	-----

Chapter 4

Table 1. Comparison of environmental conditions, parameters sampled and methodological information from three studies used in this meta-analysis.	161
---	-----

Table 2. Species names and supporting phylogenetic information for modeled populations.	162
---	-----

Table 3. Percent of total isolates in each population associated with large ($> 64 \mu\text{m}$) particles from both fract-2006 and particle-2007	164
---	-----

Table 4. List of sequences used in multi-locus gene sequence	165
--	-----

assembly using unassembled Illumina sequenced fragments.

Table 5. Accession numbers for *V. breoganii* sequences used in this analysis, which were not included in previous phylogenetic analyses.

172

Chapter One

Introduction

INTRODUCTION

This thesis investigates key aspects of bacterial ecology and population biology using *Vibrionaceae* in a coastal ecosystem as a model. Bacterial population structure is the distribution and assembly of populations in space and time. Many mechanisms that generate or maintain genetic diversity, such as speciation, dispersal and species interactions, can be brought to light by investigating population structure (Martiny *et al.*, 2006). Building on previous work facilitating the identification of ecologically relevant bacterial populations, this thesis uses the association of bacterial populations with invertebrate hosts and specific types of particles in the water column to identify reproducible and specialized associations in the environment. It also investigates whether mechanisms of competitive exclusion can explain the observation. All of these efforts extend our understanding of bacterial evolution and speciation. Here, I review the current knowledge of bacterial population biology and community structure, including mechanisms that create population structure within host associated microbial populations and on particles in the water column, focusing on associations within the *Vibrionaceae*.

Identifying Bacterial Populations

The first challenge in addressing bacterial population structure is defining an appropriate measure of the bacterial population. Populations for plant and animal taxa are groups of organisms of the same species that live together and interbreed. Given that the species concept for microorganisms is mired with controversy and lacks a solid

theoretical foundation (Fraser *et al.*, 2009; Gevers *et al.*, 2005), defining a group of organisms of the same species is the first hurdle in studies of microbial population structure. Although taxonomists have tried to conform bacterial species concept to the species concept for eukaryotes, difference in the rate of recombination (Lawrence, 2002; Vos & Didelot, 2009) and extensive horizontal gene transfer between distantly related groups (Ge *et al.*, 2005) make it difficult to compare microbial and eukaryotic species. Commonly, in studies of bacterial diversity, a genetic cut-off value is used to group bacterial diversity into populations or species rather than applying a model of ecological differentiation (Acinas *et al.*, 2004; Furlong *et al.*, 2002; Huber *et al.*, 2007; Schloss & Handelsman, 2005). This proxy for bacterial species, often called operational taxonomic units (OTUs), are predicted from clone library sequences using similarity cut-off values from small subunit rRNA (16S) sequences. As presently defined bacterial species and OTUs are not chosen through principles of ecology or evolution, they may not be the best units with which to investigate population structure.

Recent advances have been made to group bacteria by ecological (Hunt *et al.*, 2008a) and evolutionary (Koeppel *et al.*, 2008) principles, thus more appropriately identifying populations that should be considered in these ecological studies. Koeppel *et al.* (2008) used a strictly evolutionary model to identify the fingerprints of selection on the phylogeny of *Bacillus* clades isolated from soil samples across a range of environments within a canyon using multiple genetic loci. Although evidence for ecological differentiation can be gathered from properties of the strains after the model is run, the actual distribution of organisms in the environment is not used directly to assemble

organisms into ecological differentiated groups. In an alternative approach, Hunt et. al (2008) provided direct evidence of ecological differentiation from the distribution of genetically similar organisms across sampled environmental categories. They use an algorithm to determine how much of the total diversity should be contained within each population by i.) conservatively estimating the number of true habitats (or niches) from the combined phylogenetic and environmental data and ii.) organizing strains into populations based on their predicted habitat and genetic similarity. While this approach may fail to detect groups that are ecologically differentiated across unsampled environmental axes, the resulting groups can be confidently related to the environmental categories that define their habitat. Applying one of these approaches to the study of microbial population structure significantly reduces the possibility of combining ecologically distinct populations, which might occur using arbitrarily defined units or species.

Population Structure and Relative Fitness

When looking at population distribution in space and time, understanding the chemical and physical characteristics of sampled areas will be crucial in bringing to light the mechanisms at work creating any observed structure. Sampled areas must be spatially separated (even if only by millimeters), but dispersal of organisms between areas and the chemical, physical and biological similarities between areas will both impact community structure. Ideally, the relative contribution of each of these factors can be discerned through careful and controlled experimental design.

While it is clear that microorganisms are not distributed evenly in space, one commonly held view is that they are distributed efficiently, but local environmental conditions select for specific populations to become relatively more abundant. In situations where this may be a good approximation of the primary mechanisms creating population structure, one can invoke niche theory to understand how organisms interact with other species and their environment (Tilman, 1982). Under this model, competition with co-existing organisms is an important determinant of which species are present in a specific environment. Over evolutionary time scales, competition can cause species to adapt alternative life-styles, e.g., to attach to different types of particles or to utilize different substrates on the same particle. Such trade-offs are a necessary result of constraints on cellular machinery, such as cell-surface proteins and enzymatic capacity, or genome size, resulting in specialized populations with limited niche overlap.

Temperature has been shown in many systems to be an important determinant of population abundances. Species often have different relative growth rates at different temperatures (Tilman *et al.*, 1981), resulting in shifts in population relative abundances over the course of a season. Within the *Vibrionaceae*, many populations are influenced by changes in water temperature (Motes *et al.*, 1998; Ruby & Nealson, 1978) and distinct year-round and warm water populations have been observed (Hunt *et al.*, 2008a; Randa *et al.*, 2004; Thompson *et al.*, 2004c).

Environmental conditions, such as depth, nutrients and competitor abundance can also have a strong impact on the distribution of phytoplankton in the ocean. *Pelagibacter ubique*, previously known as SAR11, is one of the most abundant organisms in the ocean.

Small subunit rDNA gene clusters correspond with distinct depth profiles in the ocean, suggesting that divergence within this clade results from niche specialization (Field *et al.*, 1997). In the globally important bacterial photoautotroph, *Prochlorococcus*, distinct populations assemble according to light, nutrients and competitor abundance (Johnson *et al.*, 2006; Rocap *et al.*, 2003). A good approximation of the global distribution of all phytoplankton types emerges from an ocean circulation model initialized randomly with cells representing all different phytoplankton types and essentially allowing the environment to select for abundant types according to competition (Follows *et al.*, 2007). This evidence suggests that, especially in the ocean, environmental conditions can have a noticeable impact on the assembly of bacterial populations.

Population Structure and Dispersal

There is evidence that dispersal, even in microorganisms, may play a role in shaping population structure. Limited dispersal, across continents for example, has been demonstrated to influence speciation in plant and animal taxa (Anderson, 1994). While we are far from understanding how limited dispersal effects speciation in microorganisms, it is clear that distance can have an affect on bacterial community structure. Evidence of the effects of geographic separation has been shown for archaeal inhabitants of hot springs, *Sulfolobous*, whose habitats are separated by large stretches of uninhabitable environments (Whitaker *et al.*, 2003). Geographic separation was also shown to be important in cyanobacterial structure across hot springs (Papke *et al.*, 2003). Effects of limited dispersal may be particularly pronounced for specific types of

microorganisms, such as those that do not form spores or resistant forms when removed from ideal growth conditions, or microorganisms with specific and uncommon growth requirements. However, the ocean seems to be an area where geographic isolation is uncommon. Studies of the diatoms fossil record revealed that geographic separation of these eukaryotic microbes is not maintained for long periods of time (Cermenó & Falkowski, 2009). In less well-mixed environments, such as soil samples, the effects of distance on bacterial populations seem to be negligible if samples are within a few kilometers (Martiny *et al.*, 2006). It remains to be seen how the effects of geographical isolation may affect speciation of all types of microorganisms from a wide range of habitat types.

Colonization dynamics can also contribute substantially to observed population structure if only a small portion of all populations that can potentially grow initially colonize an area. As may be the case in habitats that are created and turned over at a relatively high rate, population structure across areas of concentrated nutrients (patches) could be the result of population bottlenecks during initial colonization more than specific growth effects within the habitat (Fig. 1) (Fraser *et al.*, 2009). This structure may describe the situation where bacteria use limited resource intensively for a short time, and are then randomly dispersed to other patches. Selective sweeps are expected to reduce diversity in a simple, well mixed environment, but not when populations are highly divided (Fraser *et al.*, 2009). The resulting meta-population structure is due to stochastic factors, assuming multiple populations can grow equally well and colonization determines the allele frequency within the patch. Gut microbiota are a unique assembly of

populations (Eckburg *et al.*, 2005; Gao *et al.*, 2007; Pei *et al.*, 2004) drawn from a diverse group of highly adapted bacteria (Dethlefsen *et al.*, 2006; Ley *et al.*, 2006). It remains to be seen whether populations within an individual are uniquely adapted to the highly individualized conditions in each person, or are formed through stochastic mechanisms of colonization. Stochastic colonization by only a few bacterial populations that creates population structure can be distinguished from succession (Connell & Slatyer, 1977), if the age of the habitat can be determined. Habitats of the same age would have the same population(s) if succession occurs through a defined progression of population abundances, but are not expected to be similar if a stochastic association with only a few of multiple possible populations occurs.

Population structure may arise or be influenced by high relative migration rates of specific populations into the sampled area. The resulting population structure would be (partially) uncoupled from specific growth because constant migration might alter the dynamics. The presence of thermophilic *Firmicutes* on the cold ocean sea-floor could only be explained by passive dispersal to the area since these organisms were not metabolically active at *in situ* temperatures (Hubert *et al.*, 2009). The abundance of these thermophiles on the cold sea floor, while low, was not related to how well the organism grows within that area, but instead the rate of migration presumably resulting from advective flows of fluids from hundreds of meters below the seafloor. Many of the techniques used in microbial ecology do not differentiate between cells that are dormant and actively growing, and detected organisms could be in transit. The difficulty of

distinguishing between present and active cells for studies of microbiology might make relative migration rates an important aspect of microbial community structure.

Population Structure of Particle Attached Bacteria

Particles represent an important resource for bacteria, since most are colonized (Long & Azam, 1996) and up to half of all bacteria present in the water may be attached to particles (Crump *et al.*, 1998) depending on location. Thus, competition for this abundant nutrient source may play as strong role in structuring bacterial communities. Particles, operationally defined as being retained on a filter of approximately 0.45 μm are comprised of a wide range of chemically and physically distinct forms (Simon *et al.*, 2002). Particulate organic matter is mainly derived from living and dead bacteria, algae, and zooplankton (Bishop *et al.*, 1977; Gibbs, 1974). The wide variety of forms invites adaptation of bacteria to attach to particles of different types. Many particles are quite ephemeral and their composition changes as decomposition progresses, opening up the possibility of bacterial specialization for certain stage of succession. For example, by approximately three days at 22 °C, copepod remain are devoid of internal structure and consist exclusively of the chitinous carapace (Tang *et al.*, 2006). As particulate matter becomes depleted of key nutrients, such as nitrogen and phosphorous, the composition of bacteria attached to these particles could change (Jacobsen & Azam, 1984). Bacterial specialization for either specific types of particles, or different stages of particle degradation when different nutrients are present, may occur. However, detecting this specialization, especially using population structure, could be problematic because of the

diversity of particle types, ephemeral nature of particles (different stages of degradation).

Partitioning of bacterial populations in the water column may be influenced by a combination of both specific growth/survival and migration. Most studies have found differences between bacterial populations attached to particles in the water and those found free-living (Acinas *et al.*, 1999; Crump *et al.*, 1999; Delong *et al.*, 1993). Others have found little difference, possibly since the particle attached fraction likely de-attach to form part of the free-living fraction (Abell & Bowman, 2005). Only a portion of the total bacterial population attached to particles was shown to be active (Ghiglione *et al.*, 2009), thus some occupants may take advantage of the benefits of particle attachment without active metabolism. Particle size preferences were demonstrated for *Vibrio* populations, which were found preferentially in the free-living or attached fraction of seawater (Hunt *et al.*, 2008a). This may represent different patterns of succession on particles, since the smaller particles are possibly derived from larger particles and would likely be depleted in many important nutrients, or adaptation to specific particle types since some particles are expected to be unique (i.e. comprised of zooplankton and phytoplankton that never reach the size of the larger fraction).

Attachment to zooplankton has been demonstrated as an important aspect of *Vibrio* biology and public health. Zooplankton tend to harbor more vibrios than phytoplankton (Simidu *et al.*, 1977). The number of bacteria found internally in zooplankton is relatively consistent, but external bacterial counts can vary with water temperature (Kaneko & Colwell, 1978). Attachment of *V. cholerae*, the causative agent of cholera, is site specific, usually to mouth parts and egg casings (Huq *et al.*, 2005; Huq

et al., 1983) Once attached, they increase in number. Their presence on zooplankton has been suggested to be an important aspect of the transmittance of cholera (Faruque *et al.*, 2006).

Microbiota of Animals

Bacteria are attached to nearly all animal surfaces that are exposed to the environment and play important roles in the health of the organism. Bacteria are particularly abundant in the gut, where nutrients are concentrated and conditions are fairly stable. Associations between animals and gut microorganisms are ancient, as demonstrated by the coevolution of microbes in a wide range mammalian species (Ley *et al.*, 2008). These relationship have been developed from the additional metabolic benefits associated microorganisms can provide to its host, as demonstrated by linking microbiome and metabolic phenotype variation (Li *et al.*, 2008). Thus, while pathogenic bacteria have received attention as the agents of disease, most of the bacterial associations with humans may actually be necessary for proper health.

Ancient association and heterogeneity in physical and chemical properties along the gut suggest that different bacterial groups might have highly specialized functions within the community. Animals represent highly selective environments since only a small subset of microbial phyla can associate with them. In human guts, for example, only 2 of 55 bacterial and 1 of 13 archaeal divisions occur in high proportions leading to the suggestion that the host selects according to characteristics that are stable over evolutionary time and multiple individuals (Dethlefsen *et al.*, 2007; Ley *et al.*, 2006).

However, when analyzed at fine-scale phylogenetic resolution, individual hosts may harbor a fairly unique assembly of microbes (Eckburg *et al.*, 2005; Gao *et al.*, 2007; Pei *et al.*, 2004). It is unknown whether the extensive variation seen at the individual level is the result of bacterial responses to immunological and dietary difference, or if stochastic associations are responsible for such differences. Microbial communities may assemble from populations of specialist strains, which are adapted to the specific conditions presented by a host individual; alternatively, microbial communities may draw from pools of generalists, distinguished largely by their ability to thrive within (certain types of) animals.

GOALS OF THIS THESIS

This thesis investigates whether populations of the gamma-proteobacterial family, *Vibrionaceae*, display highly specialized distribution across a wide range of environmental categories, such as marine invertebrates or particles in the water column. Additionally, it seeks to determine whether *in situ* population distributions can be attributable to competition between other vibrio populations, through *in vitro* competition assays on native particles. Finally, populations identified in this thesis are compared to those observed in previous studies to determine whether populations are consistently predicted from associations with different environmental categories or correspond to

named species. These experiments highlight processes that shape and maintain diversity in natural bacterial populations.

The members of the *Vibrio* genus were used as a model system because they are an important part of the bacterioplankton community and extensive diversity has been documented in this group (Thompson *et al.*, 2005b), making them amenable for use in my study. Many members are pathogenic to both humans and marine biota, and three significant human pathogens in this genus have been identified (Thompson *et al.*, 2004a). *Vibrio* species are highly abundant in many coastal environments and are important in nutrient cycling, especially as degraders of the highly abundant marine polymer, chitin. Members of the *Vibrio* genus have been found at high densities in diverse environments within the water column of the coastal ocean. Studies have determined that vibrios can be found in microaggregates (Simon *et al.*, 2002), biofilms (Kwon *et al.*, 2002), sediments (Kopke *et al.*, 2005), associated with marine animals and macroalgae (Sawabe *et al.*, 2004; Thomson *et al.*, 2005) and within the water column. Vibrios are amenable to multilocus sequencing because they can be isolated on selective media from the environment and grown in culture. Thus, the *Vibrio* genus is a tractable model system to use in this study.

The environmental distribution of vibrios across marine invertebrates (Chapter 2), demonstrated that *Vibrio* populations show little specialization for specific host types or host tissue. The preponderance of generalist populations suggests that migration into invertebrate hosts is high and corresponding selective pressure may be low. Population structure between host individuals likely created by clonal expansion was significant in

crab and zooplankton samples, while little structure was observed between mussel individuals. This suggests that metapopulation structure, where the populations is subdivided into patches and individuals can disperse between patches, is important for bacteria colonizing animals. Within the water column, bacterial population specialization for particles of different types was predicted from *in situ* distribution across plant-derived and zooplankton samples (Chapter 3). Relative fitness between specialized populations was tested on native particles *in vitro* to determine the effects of competition in generating population structure. Extensive resource partitioning of both particle types and particle specific substrates was demonstrated, suggesting competitive exclusion as an important factor shaping distributions. Yet, competition experiments of specialized populations on native particles could not explain *in situ* distributions without invoking difference in migration. Additionally, the reproducibility and specificity of *Vibrio* population associations was investigated across years by comparing a previous study of populations associated with water column particles to the associations determined from Chapters 2 and 3 of this thesis (Chapter 4). Reproducibility across years was high, but not all associations were reproduced. Specificity, as defined by consistently limited distribution across all environmental categories, was low, since many populations were found in almost all of the sampled categories. Finally, bacterial populations predicted from environmental associations were largely consistent with named *Vibrio* species. This information can be used to refine our current thinking of the species concept in microbiology. All of this work highlights the importance affects of migration and competition on the diversity of co-occurring *Vibrio* populations in the coastal ocean.

REFERENCES

- Abell, G.C.J. & Bowman, J.P. (2005) Colonization and community dynamics of class Flavobacteria on diatom detritus in experimental mesocosms based on Southern Ocean seawater. *Fems Microbiology Ecology*, **53**, 379-391.
- Acinas, S.G., Anton, J., & Rodriguez-Valera, F. (1999) Diversity of free-living and attached bacteria in offshore western Mediterranean waters as depicted by analysis of genes encoding 16S rRNA. *Applied and Environmental Microbiology*, **65**, 514-522.
- Acinas, S.G., Klepac-Ceraj, V., Hunt, D.E., Pharino, C., Ceraj, I., Distel, D.L., & Polz, M.F. (2004) Fine-scale phylogenetic architecture of a complex bacterial community. *Nature*, **430**, 551-554.
- Anderson, S. (1994) Area and Endemism. *Quarterly Review of Biology*, **69**, 451-471.
- Bishop, J.K.B., Edmond, J.M., Ketten, D.R., Bacon, M.P., & Silker, W.B. (1977) Chemistry, Biology, and Vertical Flux of Particulate Matter from Upper 400 M of Equatorial Atlantic Ocean. *Deep-Sea Research*, **24**, 511-&.
- Cermeno, P. & Falkowski, P.G. (2009) Controls on Diatom Biogeography in the Ocean. *Science*, **325**, 1539-1541.
- Connell, J.H. & Slatyer, R.O. (1977) Mechanisms of Succession in Natural Communities and Their Role in Community Stability and Organization. *American Naturalist*, **111**, 1119-1144.
- Crump, B.C., Armbrust, E.V., & Baross, J.A. (1999) Phylogenetic analysis of particle-attached and free-living bacterial communities in the Columbia river, its estuary, and the adjacent coastal ocean. *Applied and Environmental Microbiology*, **65**, 3192-3204.
- Crump, B.C., Baross, J.A., & Simenstad, C.A. (1998) Dominance of particle-attached bacteria in the Columbia River estuary, USA. *Aquatic Microbial Ecology*, **14**, 7-18.
- DeLong, E.F., Franks, D.G., & Alldredge, A.L. (1993) Phylogenetic Diversity of Aggregate-Attached Vs Free-Living Marine Bacterial Assemblages. *Limnology and Oceanography*, **38**, 924-934.
- Dethlefsen, L., Eckburg, P.B., Bik, E.M., & Relman, D.A. (2006) Assembly of the human intestinal microbiota. *Trends in Ecology & Evolution*, **21**, 517-523.
- Dethlefsen, L., McFall-Ngai, M., & Relman, D.A. (2007) An ecological and evolutionary perspective on human-microbe mutualism and disease. *Nature*, **449**, 811-818.
- Eckburg, P.B., Bik, E.M., Bernstein, C.N., Purdom, E., Dethlefsen, L., Sargent, M., Gill, S.R., Nelson, K.E., & Relman, D.A. (2005) Diversity of the human intestinal microbial flora. *Science*, **308**, 1635-1638.
- Faruque, S.M., Biswas, K., Udden, S.M.N., Ahmad, Q.S., Sack, D.A., Nair, G.B., & Mekalanos, J.J. (2006) Transmissibility of cholera: In vivo-formed biofilms and their relationship to infectivity and persistence in the environment. *Proceedings of*

- the National Academy of Sciences of the United States of America*, **103**, 6350-6355.
- Field, K.G., Gordon, D., Wright, T., Rappe, M., Urbach, E., Vergin, K., & Giovannoni, S.J. (1997) Diversity and depth-specific distribution of SAR11 cluster rRNA genes from marine planktonic bacteria. *Applied and Environmental Microbiology*, **63**, 63-70.
- Follows, M.J., Dutkiewicz, S., Grant, S., & Chisholm, S.W. (2007) Emergent biogeography of microbial communities in a model ocean. *Science*, **315**, 1843-1846.
- Fraser, C., Alm, E.J., Polz, M.F., Spratt, B.G., & Hanage, W.P. (2009) The Bacterial Species Challenge: Making Sense of Genetic and Ecological Diversity. *Science*, **323**, 741-746.
- Furlong, M.A., Singleton, D.R., Coleman, D.C., & Whitman, W.B. (2002) Molecular and culture-based analyses of prokaryotic communities from an agricultural soil and the burrows and casts of the earthworm *Lumbricus rubellus*. *Applied and Environmental Microbiology*, **68**, 1265-1279.
- Gao, Z., Tseng, C.H., Pei, Z.H., & Blaser, M.J. (2007) Molecular analysis of human forearm superficial skin bacterial biota. *Proceedings of the National Academy of Sciences of the United States of America*, **104**, 2927-2932.
- Ge, F., Wang, L.S., & Kim, J. (2005) The cobweb of life revealed by genome-scale estimates of horizontal gene transfer. *Plos Biology*, **3**, 1709-1718.
- Gevers, D., Cohan, F.M., Lawrence, J.G., Spratt, B.G., Coenye, T., Feil, E.J., Stackebrandt, E., Van de Peer, Y., Vandamme, P., Thompson, F.L., & Swings, J. (2005) Re-evaluating prokaryotic species. *Nature Reviews Microbiology*, **3**, 733-739.
- Ghiglione, J.F., Conan, P., & Pujo-Pay, M. (2009) Diversity of total and active free-living vs. particle-attached bacteria in the euphotic zone of the NW Mediterranean Sea. *Fems Microbiology Letters*, **299**, 9-21.
- Gibbs, R.J. (1974) *Suspended solids in water* Pleum Press, New York.
- Huber, J.A., Mark Welch, D., Morrison, H.G., Huse, S.M., Neal, P.R., Butterfield, D.A., & Sogin, M.L. (2007) Microbial population structures in the deep marine biosphere. *Science*, **318**, 97-100.
- Hubert, C., Loy, A., Nickel, M., Arnosti, C., Baranyi, C., Bruchert, V., Ferdelman, T., Finster, K., Christensen, F.M., de Rezende, J.R., Vandieken, V., & Jorgensen, B.B. (2009) A Constant Flux of Diverse Thermophilic Bacteria into the Cold Arctic Seabed. *Science*, **325**, 1541-1544.
- Hunt, D.E., David, L.A., Gevers, D., Preheim, S.P., Alm, E.J., & Polz, M.F. (2008) Resource partitioning and sympatric differentiation among closely related bacterioplankton. *Science*, **320**, 1081-1085.
- Huq, A., Sack, R.B., Nizam, A., Longini, I.M., Nair, G.B., Ali, A., Morris, J.G., Khan, M.N.H., Siddique, A.K., Yunus, M., Albert, M.J., Sack, D.A., & Colwell, R.R. (2005) Critical factors influencing the occurrence of *Vibrio cholerae* in the environment of Bangladesh. *Applied and Environmental Microbiology*, **71**, 4645-4654.

- Huq, A., Small, E.B., West, P.A., Huq, M.I., Rahman, R., & Colwell, R.R. (1983) Ecological Relationships between *Vibrio-Cholerae* and Planktonic Crustacean Copepods. *Applied and Environmental Microbiology*, **45**, 275-283.
- Jacobsen, T.R. & Azam, F. (1984) Role of Bacteria in Copepod Fecal Pellet Decomposition - Colonization, Growth-Rates and Mineralization. *Bulletin of Marine Science*, **35**, 495-502.
- Johnson, Z.I., Zinser, E.R., Coe, A., McNulty, N.P., Woodward, E.M.S., & Chisholm, S.W. (2006) Niche partitioning among *Prochlorococcus* ecotypes along ocean-scale environmental gradients. *Science*, **311**, 1737-1740.
- Kaneko, T. & Colwell, R.R. (1978) Annual Cycle of *Vibrio-Parahaemolyticus* in Chesapeake Bay. *Microbial Ecology*, **4**, 135-155.
- Koeppel, A., Perry, E.B., Sikorski, J., Krizanc, D., Warner, A., Ward, D.M., Rooney, A.P., Brambilla, E., Connor, N., Ratcliff, R.M., Nevo, E., & Cohan, F.M. (2008) Identifying the fundamental units of bacterial diversity: A paradigm shift to incorporate ecology into bacterial systematics. *Proceedings of the National Academy of Sciences*, **105**, 2504-2509.
- Kopke, B., Wilms, R., Engelen, B., Cypionka, H., & Sass, H. (2005) Microbial diversity in coastal subsurface sediments: a cultivation approach using various electron acceptors and substrate gradients. *Applied and Environmental Microbiology*, **71**, 7819-7830.
- Kwon, K.K., Lee, H.S., Jung, S.Y., Yim, J.H., Lee, J.H., & Lee, H.K. (2002) Isolation and identification of biofilm-forming marine bacteria on glass surfaces in Dae-Ho Dike, Korea. *Journal of Microbiology*, **40**, 260-266.
- Lawrence, J.G. (2002) Gene transfer in bacteria: Speciation without species? *Theoretical Population Biology*, **61**, 449-460.
- Ley, R.E., Hamady, M., Lozupone, C., Turnbaugh, P.J., Ramey, R.R., Bircher, J.S., Schlegel, M.L., Tucker, T.A., Schrenzel, M.D., Knight, R., & Gordon, J.I. (2008) Evolution of mammals and their gut microbes. *Science*, **320**, 1647-1651.
- Ley, R.E., Peterson, D.A., & Gordon, J.I. (2006) Ecological and evolutionary forces shaping microbial diversity in the human intestine. *Cell*, **124**, 837-848.
- Li, M., Wang, B.H., Zhang, M.H., Rantalainen, M., Wang, S.Y., Zhou, H.K., Zhang, Y., Shen, J., Pang, X.Y., Zhang, M.L., Wei, H., Chen, Y., Lu, H.F., Zuo, J., Su, M.M., Qiu, Y.P., Jia, W., Xiao, C.N., Smith, L.M., Yang, S.L., Holmes, E., Tang, H.R., Zhao, G.P., Nicholson, J.K., Li, L.J., & Zhao, L.P. (2008) Symbiotic gut microbes modulate human metabolic phenotypes. *Proceedings of the National Academy of Sciences of the United States of America*, **105**, 2117-2122.
- Long, R.A. & Azam, F. (1996) Abundant protein-containing particles in the sea. *Aquatic Microbial Ecology*, **10**, 213-221.
- Martiny, J.B.H., Bohannan, B.J.M., Brown, J.H., Colwell, R.K., Fuhrman, J.A., Green, J.L., Horner-Devine, M.C., Kane, M., Krumins, J.A., Kuske, C.R., Morin, P.J., Naeem, S., Ovrees, L., Reysenbach, A.L., Smith, V.H., & Staley, J.T. (2006) Microbial biogeography: putting microorganisms on the map. *Nature Reviews Microbiology*, **4**, 102-112.

- Motes, M.L., DePaola, A., Cook, D.W., Veazey, J.E., Hunsucker, J.C., Garthright, W.E., Blodgett, R.J., & Chirtel, S.J. (1998) Influence of water temperature and salinity on *Vibrio vulnificus* in northern Gulf and Atlantic Coast oysters (*Crassostrea virginica*). *Applied and Environmental Microbiology*, **64**, 1459-1465.
- Papke, R.T., Ramsing, N.B., Bateson, M.M., & Ward, D.M. (2003) Geographical isolation in hot spring cyanobacteria. *Environmental Microbiology*, **5**, 650-659.
- Pei, Z.H., Bini, E.J., Yang, L.Y., Zhou, M.S., Francois, F., & Blaser, M.J. (2004) Bacterial biota in the human distal esophagus. *Proceedings of the National Academy of Sciences of the United States of America*, **101**, 4250-4255.
- Randa, M.A., Polz, M.F., & Lim, E. (2004) Effects of temperature and salinity on *Vibrio vulnificus* population dynamics as assessed by quantitative PCR. *Applied and Environmental Microbiology*, **70**, 5469-5476.
- Rocap, G., Larimer, F.W., Lamerdin, J., Malfatti, S., Chain, P., Ahlgren, N.A., Arellano, A., Coleman, M., Hauser, L., Hess, W.R., Johnson, Z.I., Land, M., Lindell, D., Post, A.F., Regala, W., Shah, M., Shaw, S.L., Steglich, C., Sullivan, M.B., Ting, C.S., Tolonen, A., Webb, E.A., Zinser, E.R., & Chisholm, S.W. (2003) Genome divergence in two *Prochlorococcus* ecotypes reflects oceanic niche differentiation. *Nature*, **424**, 1042-1047.
- Ruby, E.G. & Neelson, K.H. (1978) Seasonal-Changes in Species Composition of Luminous Bacteria in Nearshore Seawater. *Limnology and Oceanography*, **23**, 530-533.
- Sawabe, T., Hayashi, K., Moriwaki, J., Thompson, F.L., Swings, J., Potin, P., Christen, R., & Ezura, Y. (2004) *Vibrio gallicus* sp nov., isolated from the gut of the French abalone *Haliotis tuberculata*. *International Journal of Systematic and Evolutionary Microbiology*, **54**, 843-846.
- Schloss, P.D. & Handelsman, J. (2005) Introducing DOTUR, a computer program for defining operational taxonomic units and estimating species richness. *Applied and Environmental Microbiology*, **71**, 1501-1506.
- Simidu, U., Kaneko, E., & Taga, N. (1977) Microbiological Studies of Tokyo Bay. *Microbial Ecology*, **3**, 173-191.
- Simon, M., Grossart, H.P., Schweitzer, B., & Ploug, H. (2002) Microbial ecology of organic aggregates in aquatic ecosystems. *Aquatic Microbial Ecology*, **28**, 175-211.
- Tang, K.W., Freund, C.S., & Schweitzer, C.L. (2006) Occurrence of copepod carcasses in the lower Chesapeake Bay and their decomposition by ambient microbes. *Estuarine Coastal and Shelf Science*, **68**, 499-508.
- Thompson, F.L., Iida, T., & Swings, J. (2004a) Biodiversity of vibrios. *Microbiology and Molecular Biology Reviews*, **68**, 403-+.
- Thompson, J.R., Pacocha, S., Pharino, C., Klepac-Ceraj, V., Hunt, D.E., Benoit, J., Sarma-Rupavtarm, R., Distel, D.L., & Polz, M.F. (2005) Genotypic diversity within a natural coastal bacterioplankton population. *Science*, **307**, 1311-1313.
- Thompson, J.R., Randa, M.A., Marcelino, L.A., Tomita-Mitchell, A., Lim, E., & Polz, M.F. (2004b) Diversity and dynamics of a north Atlantic coastal *Vibrio* community. *Applied and Environmental Microbiology*, **70**, 4103-4110.

- Thomson, R., Macpherson, H.L., Riaza, A., & Birkbeck, T.H. (2005) *Vibrio splendidus* biotype 1 as a cause of mortalities in hatchery-reared larval turbot, *Scophthalmus maximus* (L.). *Journal of Applied Microbiology*, **99**, 243-250.
- Tilman, D. (1982) *Resource Competition and Community Structure* Princetown University Press, Princetown.
- Tilman, D., Mattson, M., & Langer, S. (1981) Competition and Nutrient Kinetics Along a Temperature-Gradient - an Experimental Test of a Mechanistic Approach to Niche Theory. *Limnology and Oceanography*, **26**, 1020-1033.
- Vos, M. & Didelot, X. (2009) A comparison of homologous recombination rates in bacteria and archaea. *Isme Journal*, **3**, 199-208.
- Whitaker, R.J., Grogan, D.W., & Taylor, J.W. (2003) Geographic barriers isolate endemic populations of hyperthermophilic archaea. *Science*, **301**, 976-978.

FIGURE

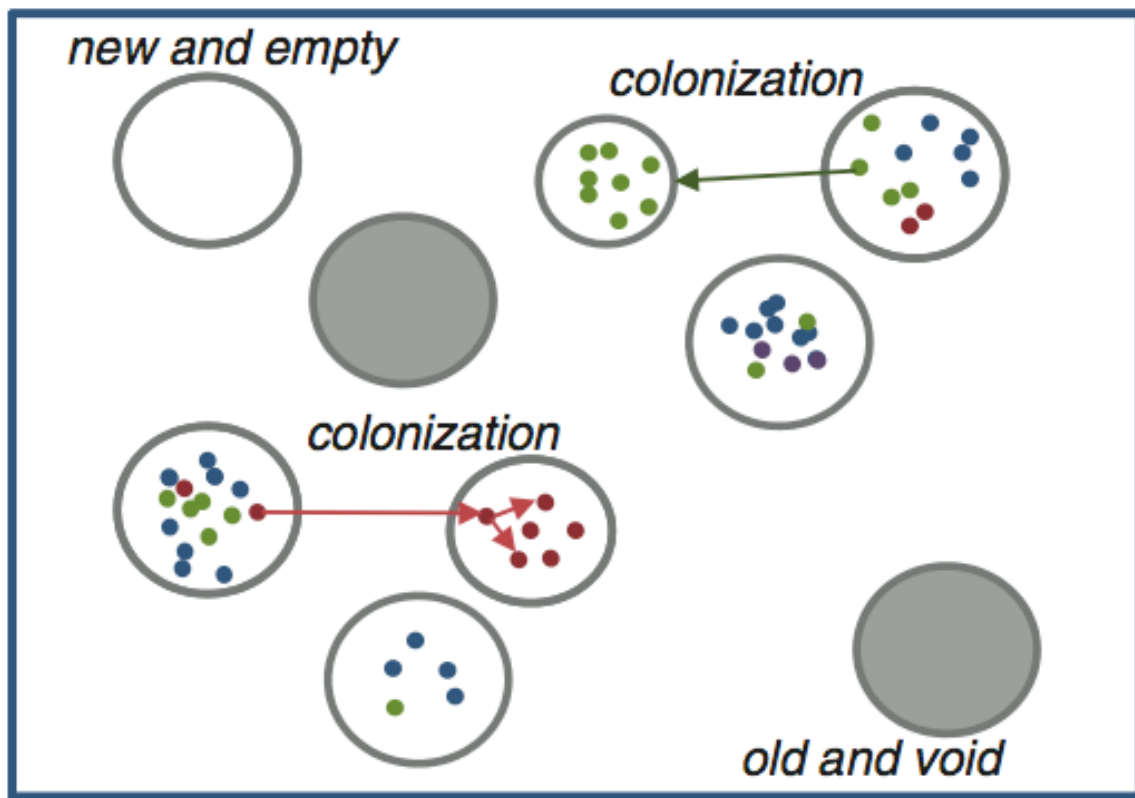


Fig. 1 Model of bacterial evolution resulting in low effective population sizes (bottleneck). Patches of varying size (gray circles) are vacant (empty) or may be colonized by a single genotype randomly acquired from another patch. Strains may diversify within a patch (as shown by different colors representing distinct genotypes), which may colonize empty patches as described above. A characteristic of this sort of metapopulation is patch turnover, in which patches occasionally become unable to support colonization and their inhabitants are removed (solid gray circles). From Fraser et al (2009).

Chapter Two

Generalists Dominate *Vibrio* Populations in the Microbiome of Marine Animals¹

¹ To be submitted with Yan Boucher, Hans Wildschutte, Lawrence A. David, Daniele Veneziano, Eric J. Alm, Martin F. Polz

ABSTRACT

Although the microbiome is increasingly recognized to influence animal health, the extent to which animals represent highly selective habitats for microbes leading to predominance of high host specificity remains poorly understood. Here, we show that vibrios, which are well-known commensals and opportunistic pathogens of marine animals, overall display little host preference, likely because of efficient dispersal-colonization dynamics mediated by food items. We isolated 1,753 strains from water and animal samples, which are linked in a food chain and display different degrees of habitat similarity (respiratory and digestive tract of mussels and crabs, live and dead zooplankton, and whole water samples). Multilocus sequence data served as input for modeling and statistical analysis of spatio-temporal population structure. This showed that the majority of populations occurred broadly within and among hosts, with one population being a near perfect generalist with regard to seasons, host taxa and body regions. Zooplankton harbored the fewest and most specific populations, while crabs and mussels contained the highest diversity with little evidence for host preferences. Most mussel and crab associated populations were detected in water samples at similar frequencies, particularly in filter-feeding mussels where populations were also evenly distributed across host individuals. The higher variation among individuals observed in crabs and zooplankton is consistent with stochastic clonal expansions. These patterns suggest that evolution of high host specificity is rare for vibrios even though these animals represent long-lived habitats, and vibrios are consistent members of their

microbiome. Instead, the marine animal microbiome may be a fairly open system in which high rates of immigration can outpace selection for specialization.

INTRODUCTION

Animals represent complex and dynamic habitats for microbial communities, whose composition can have a profound influence on the health and nutrition of the host. Yet the ecological processes governing microbial community assembly on and in animals remain poorly understood. Recent meta-analyses of microbiome diversity on a broad phylogenetic scale have suggested that animals represent highly selective environments since only a small subset of microbial phyla can associate with them. In human guts, for example, only 2 of 55 bacterial and 1 of 13 archaeal divisions occur in high proportions leading to the suggestion that the host selects according to characteristics that are stable over evolutionary time and multiple individuals (Dethlefsen *et al.*, 2007; Ley *et al.*, 2006). On the other hand, when analyzed at fine-scale phylogenetic resolution, individual hosts may harbor a fairly unique assembly of microbes (3-5). From a microbial perspective, animals present variable environments due to individual differences in physiology, immune response and food preferences. It is therefore possible that microbiomes assemble from populations of specialist strains, which are adapted to the specific conditions presented by a host individual; alternatively, microbiomes draw from pools of generalists, distinguished largely by their ability to thrive within (certain types of) animals. In the later case, population demographics (e.g., strength of dispersal, founder effects) may have strong influence on strain composition within a host specimen.

Only few microbial populations have been mapped with sufficient phylogenetic and ecological resolution to determine whether they pursue specialist or generalist lifestyles with respect to host species and body regions. Yet such information is crucial in judging the stability and resilience of microbiomes on ecological time-scales, and how gene flow and selection shape genomic diversity within populations on evolutionary time-scales. Although little detail is available, some human associated groups of bacteria might be categorized as specialists or generalists. For example, *Escherichia coli* populations as a whole appear to be versatile, occurring in many different warm-blooded animals (Gordon & Cowling, 2003), while specific pathogenic strains can engage in transient, tissue-specific clonal expansions. Other bacterial groups are always highly tissue specific within a single host species. For example, *Neisseria gonorrhoeae* and *N. meningitides* preferentially colonize the urogenital tract and nasopharynx, respectively (Maiden, 2008). Why some populations might evolve into specialists while others retain the capability to colonize a wide variety of animal associated habitats remains poorly understood. On the one hand, evolution of specialists might be expected because animals represent stable habitats compared to the lifespan of colonizers; on the other hand, generalist populations may be maintained if dispersal among patches (individual animals) is high compared to the strength of selection (Kassen, 2002). To differentiate among these possibilities comparative analysis of population distribution at high phylogenetic resolution is required.

Here, we explore the distribution of marine bacteria of the genus *Vibrio* on fine phylogenetic and spatial scales, looking at the distribution of isolates across several

invertebrate hosts by multilocus sequence analysis (MLSA) and statistical modeling. Vibrios are constant members of the microbiome of marine animals, and several species contain well-known opportunistic human and animal pathogens (Thompson *et al.*, 2004b). Vibrios have been shown to grow rapidly within animals (Ruby & Morin, 1979) but many species also occur in the water column and in sediments (Thompson *et al.*, 2004b). We have recently shown that vibrios co-existing in the water column can be divided into closely related populations, which pursue different lifestyles (free-living, attached to particles or zooplankton) suggesting spatial resource partitioning in the water column (Hunt *et al.*, 2008c). This previous work affords the possibility to analyze animal-associated vibrios in the context of a metapopulation framework, i.e., considering potential overlap or differences of populations collected from spatially and temporally distinct habitats, which are connected by dispersal (Fig. 1). Overall, our analysis reveals surprising dominance of generalist vibrio populations with respect to animal species and body regions, where zooplankton-associated vibrio populations show higher degree of host specificity than those on larger animals (mussels and crabs). Because the same populations occur in the water column where they are specific for different particle types (Hunt *et al.*, 2008c), we suggest that that rapid dispersal-colonization dynamics (especially through active feeding behavior) dominate over selection for specialization.

RESULTS AND DISCUSSION

Study design

We explored to what extent several invertebrates represent distinct habitats for vibrios by comparing seasonal population structure (spring and fall) among several hosts. These habitats represent different degrees of chemical, physical and biological habitat similarity for colonizing bacteria and are connected through a food chain (Fig. 1). Our study design allows comparison of host species, body regions, and individuals for larger animals (mussels and crabs), and aggregates of live and dead specimen for small animals (zooplankton). Because microbes in the water column serve as potential inoculum for these animals, we concomitantly isolated vibrios from whole water samples.

A total of 1,753 *Vibrio* strains were isolated from whole seawater and 10 animal-derived sample categories. This includes the gastrointestinal and respiratory tract of 16 individual mussels and crabs, and 16 samples of ~50 individuals of either live or dead zooplankton where all samples were evenly split between spring and fall (Fig. 1). Estimation of host-associated populations was carried out in three steps: First, multilocus sequencing of three housekeeping genes (*adk*, *hsp60* and *mdh*) was used to construct a high-resolution phylogeny. Second, the phylogeny served as input to a previously developed mathematical model (AdaptML), which identifies populations as groups of related strains sharing a characteristic distribution among the sample categories ('habitat') (Hunt *et al.*, 2008c). Finally, the evenness of population distribution in the same category across samples (e.g., gills from each of the 8 crabs collected in the fall) was tested to assess predictability of host association.

Our rationale for using the AdaptML approach in this study arises from the following consideration. It combines phylogenetic and ecological information to infer

specialization with respect to the sampled environmental categories thus allowing for populations to be undifferentiated with respect to the sampled ecological categories. Other methods, such as the ecotype simulation, model populations based on their genetic structure and subsequently require searching for ecologically informative differences among populations (Cohan & Koeppel, 2008). This subtle difference is important in the context of this study since we are interested in ascertaining whether phylogenetically closely related groups are differentiated with respect to the environmental categories under consideration rather than modeling the evolutionary history of populations.

Influence of host type and body region on population structure

Our analysis suggests that the animal-derived *Vibrio* strains can be grouped into 16 populations (Fig. 2a), whose distributions among the sample categories (Fig. 2b) are summarized as four characteristic ‘habitats’ (Fig. 2c, see SI). Fifteen populations were predicted directly by the AdaptML algorithm to be differentially distributed among one or more of the 10 invertebrate and 2 seasonal sample categories. Although this suggests at least some degree of specialization, inspection of the habitat predictions shows considerable representation among most of the sample categories so that associations overall appear relatively weak (Fig. 2b,c). Further, although the majority of strains within one large cluster of closely related strains (#16) were associated with habitat H_{SFG} (Fig. 2c), *post hoc* significance testing showed this habitat prediction to be weakly supported. This is because the distribution of the sample categories on the phylogeny of these strains is not significantly different from random (not shown). These

considerations suggest that this group of strains displays characteristics of an ecological generalist population, including close genetic relationships (98.5% average nucleotide identity in MLSA genes) and near equal representation in all sample categories (Fig. 2b, *V. splendidus* population #16).

Season had a strong influence on population structure with the majority of populations being specific to fall samples while only two were predominantly found in the spring (Fig. 2a,b). This is consistent with previous culture-dependent and independent observations of strong seasonal shifts in abundance of *Vibrio* populations in temperate North Atlantic coastal waters (11, 13, 14). The generalist *V. splendidus* population #16 represents a special case having being recovered from spring and fall samples in this study (Fig. 2), while previously, populations phylogenetically indistinguishable from #16 (Fig. S1) predominantly occurred in the water column in spring samples (Hunt *et al.*, 2008c). This suggests that animals may serve as a reservoir for some *Vibrio* populations when their abundance in other habitats decreases.

Although most populations displayed some degree of host preference, the majority occurred in considerable proportions on each of the three types of animals suggesting that overall these hosts represent permissive habitats for vibrios (Fig. 2). Among the 11 populations for which we obtained >20 isolates (Fig. 2b), zooplankton appears to be the most selective host. Exclusion was particularly strong for *V. breoganii* (#6) for which 74 strains were obtained from mussels and crabs but not a single isolate from zooplankton. Conversely, crabs and mussels were hosts to most of the populations at high relative representation. For seven of the eleven populations represented by >20

isolates, greater than 40% of the individuals within a population came from crabs while for mussels and zooplankton only three populations had similar proportions, respectively. In summary, crabs and mussels contained the highest richness of *Vibrio* populations with crabs representing the preferred habitat while zooplankton, the other crustacean hosts assayed, harbored the lowest proportion of *Vibrio* populations.

Zooplankton-associated populations usually occurred on both live and dead specimens; however, they were present at much higher relative frequency on dead specimens or discarded carapaces as demonstrated by habitat H_{FZ} (Fig. 2c). This was especially pronounced for *Aliivibrio fischeri*-like (#3) and *Enterovibrio calviensis*-like (#5) populations (Fig. 2b) and suggests overall that most populations pursue a primarily saprophytic lifestyle, digesting either dead host tissue or the chitinaceous carapace. The near universal ability of vibrios to metabolize chitin (Hunt *et al.*, 2008b) is consistent with such ecological role. Only one population with no similarity to named vibrio species, *Vibrio sp. F10* (#7), was relatively more predominant on live zooplankton and may therefore be engaged in a more specific interaction, either as a commensal/symbiont or pathogen. In either case, this population may be adapted to host defense mechanisms, which in zooplankton include antimicrobial peptides (e.g., crustin) (de la Vega *et al.*, 2008; Shockey *et al.*, 2008) and melanization (Cerenius & Soderhall, 2004). Previously, zooplankton has been suggested to represent an important reservoir for vibrios, especially for *V. cholerae* (Heidelberg *et al.*, 2002). However, our analysis shows that it is important to differentiate living and dead specimens to clarify the role of these animals in *Vibrio*

ecology as has also been found in experimental infection of live and dead zooplankton with *V. cholerae* (Mueller *et al.*, 2007).

Surprisingly little population structure was observed when the distribution of vibrios on different body regions of mussels and crabs was compared (Fig. 2). The most pronounced differences were seen between digestive and respiratory tract samples with most populations occurring preferentially in the former. For example, *V. breoganii* (#6) appears to be a gut specialist with only ~5% of strains recovered from gills of crabs or mussels. In contrast, *Enterovibrio calviensis*-like (#5) is one of three populations associated with the gill of both crabs and mussels. Within the gut environment, populations occurred attached to gut tissues and in gut contents at similar relative frequency in both mussels and crabs, which might be due to contamination of tissue samples by the gut contents even though tissue samples were washed thoroughly. Moreover, little population structure was observed along the digestive tract of crabs (stomach and hindgut). Such lack of differentiation is surprising considering that hindgut and stomach presumably differ in many respects, such as available nutrients (before and after absorption in the midgut) and aeration. This observation of little spatial structure on the scales measured here suggests that if *Vibrio* population specialization within the gut environment occurs at all, it more likely happens by differential metabolite exploitation rather than adaptation to different body regions.

Population variation among host individuals

Population variation across host individuals can provide insights into the specificity of associations and/or mechanisms of colonization, especially when comparing body regions across different host species to potential source populations from the same environment (e.g., whole water samples) (Fig. 1). To identify populations with a statistically different distribution across individual mussels and crabs or zooplankton samples, we used a chi-square test of homogeneous proportions (Tarone, 1979), which is appropriate when only the relative proportion of populations in each sample is known. Since the majority of populations occur preferentially in the fall, we only considered the fall samples and grouped the smaller populations (1-3 and 9-15) to gain statistical power. The analysis identified 13 population-sample category (i.e., host body region or zooplankton sample) pairs with chi-squared values significantly higher than the $p = 0.001$ critical threshold (generated by bootstrap analysis) (Fig. 3).

The analysis revealed pronounced differences in population distribution across individual crab, zooplankton and mussel samples. Crab associated populations showed the most significant differences in distribution (Fig. 3a). Of nine sample categories with uneven population occupancy, eight were from the gastrointestinal tract and only one from gills (Fig. 3a). On zooplankton, three populations showed uneven distribution even though each of the samples consisted of ~50 specimens (as opposed to crab and mussel body regions, with a single individual for each sample) (Fig. 3b). It is possible that only few individuals are colonized at high densities since a previous report showed 3-order of magnitude variation in total *Vibrio* cells across zooplankton specimens (Heidelberg *et al.*, 2002). Alternatively, our samples of pooled individuals might have contained different

zooplankton species, only some of which are permissive for colonization by vibrios. In contrast to these crustacean hosts, mussels contained only a single population (*V. crassostreae* #8), which was not homogeneously distributed across hindgut samples (Fig. 3c). This population occurred at much lower frequency in the water column than in mussels while all other populations display similar frequencies (Fig. 3d). This suggests that this overall homogeneous population distribution in mussels may be caused by continuous inoculation through their filter-feeding mode of nutrition. Consistent with this notion is the observation that none of the populations were uneven among the eight whole water samples suggesting homogeneity on the scales sampled.

Uneven population occurrence within a single sample coincided in almost all cases with overabundance of strains belonging to a single sequence type (ST) (identical sequence in all 3 genes). In crabs and mussels, five and one of the populations uneven in one sample category consisted of >50% of a single ST, respectively. Even in zooplankton one of the populations was highly enriched in a single ST (43% of the total from one ST) (Fig. 2S). Overall, this suggests that these populations were subject to clonal expansions. These may be caused by specific host interactions, such as adaptation to host defenses, which may enable different clones to sweep within a host specimen. However, such adaptation should lead to dominance of the same clone across all samples in the same host specimen, yet this was only observed in two cases (same ST, two crab specimens). Alternatively, specific strains may arrive on food items and/or occasionally expand into ‘unoccupied’ habitats either from outside or within the animal host. Especially in crustacean hosts, the carapace, which covers most of the body, including the

stomach and hindgut, may be cleared from colonizers by molting, protozoan predation or even by animal grooming. Regardless of the mechanism, the overall rarity and seemingly haphazard distribution of large populations of identical STs among samples suggests that clonal expansions may be fairly transient.

Metapopulation structure of vibrios

Ecological theory predicts habitats that are long-lived and stable compared to the lifespan of colonizing species to be dominated by specialist populations because competition for resources will lead to tradeoffs (Kassen, 2002). In apparent contrast, we observe relatively low specificity of *Vibrio* populations for the animals analyzed, regardless of whether soft tissue or chitin-covered surfaces were sampled. Mussels and crabs appear particularly permissive for colonization and harbor most populations in significant proportions. Indeed, the dominant population across all animals is predicted to be a generalist with regard to the samples we have obtained (*V. splendidus* #16, Fig. 2).

Such low specificity of animal associated *Vibrio* populations is additionally surprising when compared to our previous study, which demonstrated strong resource partitioning among vibrios co-existing in the water column at the same site (Hunt *et al.*, 2008c). The vast majority of populations in both studies are congruent, including those previously identified as attached to zooplankton based on their recovery from samples collected with zooplankton nets (>63 μm) and now directly isolated from handpicked specimens. The main exception to population overlap between the two studies is *V. ordalii*, which was found almost exclusively free-living in the water column (Hunt *et al.*,

2008c) and is nearly absent from animals sampled in this study. However, in the water column, many of the populations had displayed stronger differences in their distribution among the sampled environmental categories even though that study had lower habitat resolution since recovery of strains from specific size fractions was used as a proxy for life-styles (free-living, and attached to particles or zooplankton) (Hunt *et al.*, 2008c). In particular, the phylogenetic cluster predicted to be an animal generalist (*V. splendidus* #16, Fig. 2) breaks up into multiple ecologically differentiated populations in the water column (Fig. S1). This suggests that groups of genomes within *V. splendidus* population #16 are adapted to exploitation of different types of particles in the water column but that they co-exist on animals without evidence of specialization.

What causes the overall weak spatial differentiation of populations in crabs and mussels while the same populations display stronger associations in the water column? As suggested above, populations may be metabolically rather than spatially differentiated in these animals; however, given the different feeding biology of mussels (filter feeders) and crabs (scavengers) many metabolic resources will be differentially distributed and specialized populations should mirror such distribution. On the other hand, high dispersal relative to the strength of selection has been suggested to prevent the evolution of specialization (Kassen, 2002; Urban *et al.*, 2008). High dispersal most certainly occurs in these populations since the animals are likely continuously inoculated from the large reservoir of populations in the water column and/or on food items (Fig. 1). This explains why populations in mussels most closely mirror those in the water column, including their homogeneous distribution across individuals, and is consistent with more skewed

population distributions in crab guts (Fig. 3). Crabs used in this study preferentially consume mussels but, as scavengers, will eat a variety of other foods. For example, detection of the remains of a polychaete worm in one of the crab stomachs correlated to a different *Vibrio* community composition than in most other crabs whose populations resembled those of mussels more closely (data not shown).

Difference in nutrient concentration in the water column and animals may mitigate competition, thus fostering co-existence of populations. Because the bulk of most populations exist in the water column, they will be primarily adapted to the relatively sparse resources there. In comparison, animal associated habitats are relatively nutrient rich so that populations may grow fairly unencumbered by others competing for resources (Ruby & Morin, 1979). This is consistent with the suggestion that if resources are abundant, species can co-exist even if they have strongly overlapping resource requirements (Tilman, 1982).

In conclusion, we suggest that metapopulation structure is highly relevant for bacteria colonizing animals. Microbiomes are diverse and different microbial taxa will show different population structure. In fact, differential community composition in the gut microbiome has been documented for many animals albeit on deep phylogenetic levels (1, 2, 24). However, the vibrios and many other microbes, which have alternate reservoirs and possess the ability to efficiently disperse laterally among hosts, may show relatively neutral population structure. This likely includes *E. coli*, whose four phylogenetic clades were previously shown to be widely distributed among warm-blooded animals (Gordon & Cowling, 2003). In particular, *E. coli* clade B1 can pursue

diverse lifestyles, it being a generalist with regards to animals (Gordon & Cowling, 2003) while dominating *E. coli* populations in freshwater environments (Walk *et al.*, 2007). Near neutral population structure has also been suggested for a number of pathogens either due to asymptomatic carriage (Roumagnac *et al.*, 2006) or overlapping microepidemics (Fraser *et al.*, 2005). Importantly, these considerations do not imply absence of selection in animals (Roumagnac *et al.*, 2006) but that even for steady members of the microbiome (such as vibrios in marine animals), population structure may be weak because selection is balanced by migration and/or adaptation to different environments.

MATERIALS AND METHODS

Water column and invertebrate sampling

Water samples and invertebrates were collected from Plum Island Sound Estuary, Ipswich, MA in the spring (April 23-May 3) and fall (September 24-October 4) of 2007 (Fig. 1). Seawater samples were collected at high tide in 4 L bottles from the shore and 5 μ l – 5 ml were directly plated or filtered onto 0.2 μ m Supor-200 filters (Pall) for bacterial isolation. Zooplankton was collected by filtering 100 L of seawater through a 64 μ m mesh net. Samples were rinsed three times with sterile seawater, washed into a 50 ml conical tube and kept at ambient temperature in the dark until processing ~2 hours later. Living and dead zooplankton were differentiated by eye under a dissecting microscope based on movement and 10-140 individuals of each category were picked from each 100 L concentrate. Four male green crabs (*Carcinus maenas*) and two male and two female

shore crabs (*Hemigrapsus sanguineus*) were collected in the spring and six male and two female shore crabs (*Hemigrapsus sanguineus*) were collected in the fall samples. Eight blue mussels (*Mytilus edulis*) were collected each season. All animals were washed with sterile seawater and placed in a whirl pack in a cooler until processing. For crabs, one gill, the entire stomach and ~4 cm of the hindgut (beginning with anus) were dissected. For mussels, approximately 1.5 cm² of gill and ~ 2 cm of the hindgut (including the anus) tissues were dissected. For both crabs and mussels, gastrointestinal contents were collected by flushing with 4 ml sterile seawater using a needle and syringe. To ensure that only firmly attached bacteria were collected from tissues, these were washed three times with sterile seawater after the contents were removed. Prior to plating for isolation, all zooplankton, crab and mussel tissue and gastrointestinal tract contents samples were homogenized in a tissue grinder and serially diluted (10- to 10,000-fold) in sterile seawater and filtered onto 0.2 µm Supor-200 filters (Pall).

Bacterial isolation and gene sequencing

Dilutions of all samples were plated on *Vibrio* selective marine TCBS media (BD Difco TCBS with 1% NaCl added) for bacterial strain isolation. About 20 colonies were picked from the most dilute samples with sufficient growth resulting in approximately 160 *Vibrio* isolates per sample type per season (Fig. 1). However, in spring samples, a few other bacteria groups grew on the selective media (e.g., *Shewanella* sp.) so that the total number of vibrios for those samples is lower. To purify strains, colonies were picked and

re-streaked three times alternating 1% TBS media (BD Bacto with 2% NaCl added) and marine TCBS media.

For gene sequencing, bacterial strains were grown in liquid culture for 2-3 days in 1% TSB at RT with shaking. A 10 μ l sample was treated with Lyse-N-Go (Thermo Fisher Scientific, Rockford, IL) to prepare DNA template. 16S primers 27f and 1492r were used to amplify the small subunit rDNA sequence (Lane, 1991a). Primers targeting the *adk* (Hunt *et al.*, 2008c), *hsp60* (H279 and H280) from (Goh *et al.*, 1996a) and *mdh* (forward primer: 5'- GAT CTG AGY CAT ATC CCW AC -3' and reverse primer: 5'- GCT TCW ACM ACY TCR GTA CCC G -3') were used to amplify and sequence part of the coding region for each gene. Additional primers were used if the initial amplification resulted in no product (*adk*: forward 5'- GCW CCD GGY GCR GGT AAA G-3' and reverse: 5'- TAG TRC CRT CRA AYT THA GGT -3'; *mdh*: forward 5'- GAY CTD AGY CAY ATC CCW AC -3'). All of the housekeeping genes were amplified using the following PCR conditions: 95°C for 3 min.; 30 cycles of 95°C for 30 sec, 37-45°C for 30 sec and 72°C for 1 min.; 72°C for 5 minutes (annealing temp for *hsp60*: 37°C, *adk*: 40°C; *mdh*: 45°C). Sequencing was performed at the Bay Paul Center at the Marine Biological Laboratories in Woods Hole. Automatic base calls were trimmed and manually curated using Sequencher (Gene Codes Corp., Ann Arbor, MI) and aligned using Clustalw (Jeanmougin *et al.*, 1998b) with visualization and further manual curation using MacClade (Sinauer Associates, Sunderland, MA). Sequences have been submitted to GenBank under accession nos. GQ988782 - GQ990534 (*hsp60*), GQ990535 - GQ992287 (*adk*), GQ992288 - GQ994040 (*mdh*).

Phylogenetic analysis

For population structure analysis, a maximum likelihood gene tree was constructed by concatenating all protein coding DNA sequences using PhyML v.2.4.5 (Guindon & Gascuel, 2003a). The GTR substitution model and 4 rate categories (parameters were estimated from the data) were used for the phylogenetic analysis. Strains identical at all loci and isolated from the same sample were included only once in the phylogeny to prevent AdaptML from identifying clonal expansions as likely habitats.

Population predictions

AdaptML was used to identify population structure (Hunt *et al.*, 2008c) using the strain phylogeny and environmental annotations (season, tissue type, gut contents). AdaptML uses a Hidden Markov Model to predict (i) the phylogenetic bounds of ecologically distinct populations and (ii) their habitat composition (distribution among environmental categories). The model threshold value was set at 0.05 and *E. coli* was used as outgroup, and the analysis was rerun 100 times with the same phylogenetic tree to verify the stability of the predictions. The iTOL tool was used to visualize the model output, habitat predictions and environmental categories (Letunic & Bork, 2007). Only populations identified by AdaptML and validated by an empirical significance test at a significance level of $p < 0.001$ (Hunt *et al.*, 2008c) were selected for further analysis, with the exception of one large group of closely related strains, which was manually assigned as a generalist population (#16 in Fig. 2).

Sampling categories were normalized based on sampling effort. Because the total number of zooplankton individuals per sample varied, the zooplankton samples (spring and fall) were each normalized to the average number of host specimens per sample (57.22). The total abundance (including clones) in each sampling category (season/host/compartiment) was normalized to the average number per category (85 strains per category) or per host (509 strains per host). Values normalized to average number per category were used to generate the bar graph in Fig. 2b, and values normalized to the average value per host are cited in the text.

Population distribution among host individuals

To estimate which populations differed across host specimens for each sample type, a version of the chi-square metric appropriate for relative proportions was used. Chi-square values were calculated as follows based on a test of goodness of fit of the binomial distribution (Tarone, 1979):

$$\chi^2 = \sum_{i=1}^M \frac{(x_i - \mu_i \bar{p})^2}{\mu_i \bar{p}(1 - \bar{p})}$$

where x_i is the observed number of isolates from a bacterial population in sample i , M is the total number of replicates, μ_i is the total number of bacteria observed in sample i , and $\bar{p} = (\sum x_i) / (\sum \mu_i)$. Parametric bootstrapping using \bar{p} and μ_i to generate random binomial values for x_i was repeated 5000 times to generate the 99 and 99.9 percentile values of χ^2 .

REFERENCES

1. Dethlefsen L, McFall-Ngai M, & Relman DA (2007) *Nature* **449**, 811-818.
2. Ley RE, Peterson DA, & Gordon JI (2006) *Cell* **124**, 837-848.
3. Eckburg PB, Bik EM, Bernstein CN, Purdom E, Dethlefsen L, Sargent M, Gill SR, Nelson KE, & Relman DA (2005) *Science* **308**, 1635-1638.
4. Gao Z, Tseng C-H, Pei Z, & Blaser MJ (2007) *Proc Natl Acad Sci U S A* **104**, 2927-2932.
5. Pei Z, Bini EJ, Yang L, Zhou M, Francois F, & Blaser MJ (2004) *Proc Natl Acad Sci U S A* **101**, 4250-4255.
6. Gordon DM & Cowling A (2003) *Microbiology* **149**, 3575-3586.
7. Maiden MCJ (2008) *Current Opinion in Microbiology* **11**, 467-471.
8. Kassen R (2002) *J. Evol. Biol.* **15**, 173-190.
9. Thompson FL, Iida T, & Swings J (2004) *Micrbiol. Mol. Biol. Rev.* **68**, 403-431.
10. Ruby EG & Morin JG (1979) *Appl Environ Microbiol* **38**, 406-411.
11. Hunt DE, Lawrence AD, Gevers D, Preheim SP, Alm EJ, & Polz MF (2008) *Science* **320**, 1081-1085.
12. Cohan FM & Koeppel AF (2008) *Curr. Biol.* **18**, R1024-R1034.
13. Thompson JR, Randa MA, Marcelino LA, Tomita A, Lim EL, & Polz MF (2004) *Appl. Environ. Microbiol.* **70**, 4103-4110.
14. Thompson JR, Pacocha S, Pharino C, Klepac-Ceraj V, Hunt DE, Benoit J, Sarma-Rupavtarm R, Distel DL, & Polz MF (2005) *Science* **307**, 1311-1313.

15. Hunt DE, Gevers D, Vahora NM, & Polz MF (2008) *Applied and Environmental Microbiology* **74**, 44-51.
16. de la Vega E, O'Leary NA, Shockey JE, Robalino J, Payne C, Browdy CL, Warr GW, & Gross PS (2008) *Molecular Immunology* **45**, 1916-1925.
17. Shockey JE, O'Leary NA, de la Vega E, Browdy CL, Baatz JE, & Gross PS (2008) *Dev. Comp. Immunol.* **33**, 668-673.
18. Cerenius L & Soderhall K (2004) *Immunological Reviews* **198**, 116-126.
19. Heidelberg JF, Heidelberg KB, & Colwell RR (2002) *Appl. Environ. Microbiol.* **68**, 5498-5507.
20. Mueller RS, McDougald D, Cusumano D, Sodhi N, Kjelleberg S, Azam F, & Bartlett DH (2007) *J. Bacteriol.* **189**, 5348-5360.
21. Tarone RE (1979) *Biometrika* **66**, 585-590.
22. Urban MC, Leibold MA, Amarasekare P, De Meester L, Gomulkiewicz R, Hochberg ME, Klausmeier CA, Loeuille N, de Mazancourt C, Norberg J, *et al.* (2008) *Trends Ecol Evol* **23**, 311-317.
23. Tilman D (1982) *Resource Competition and Community Structure* (Princeton University Press, Princeton).
24. Ley RE, Hamady M, Lozupone C, Turnbaugh PJ, Ramey RR, Bircher JS, Schlegel ML, Tucker TA, Schrenzel MD, Knight R, *et al.* (2008) *Science* **express**.
25. Walk ST, Alm EW, Cakhoun LM, Mladonicky JM, & Whittam TS (2007) *Environ. Microbiol.* **9**, 2274-2288.

26. Roumagnac P, Weill FX, Dolecek C, Baker S, Brisse S, Chinh NT, Le TA, Acosta CJ, Farrar J, Dougan G, *et al.* (2006) *Science* **314**, 1301-1304.
27. Fraser C, Hanage WP, & Spratt BG (2005) *Proc. Natl. Acad. Sci. U S A* **102**, 1968-1973.
28. Lane DJ (1991) in *Nucleic Acid Techniques in Bacterial Systematics*, eds. Stackebrandt E & Goodfellow M (Wiley & Sons, Chichester), pp. 115-175.
29. Goh SH, Potter S, Wood JO, Hemmingsen SM, Reynolds RP, & Chow AW (1996) *Journal of Clinical Microbiology* **34**, 818-823.
30. Jeanmougin F, Thompson JD, Gouy M, Higgins DG, & Gibson TJ (1998) *Trends Biochem. Sci.* **10**, 403-405.
31. Guindon S & Gascuel O (2003) *Sytematic Biology* **52**, 696-704.
32. Letunic I & Bork P (2007) *Bioinformatics* **23**, 127-128.

FIGURES

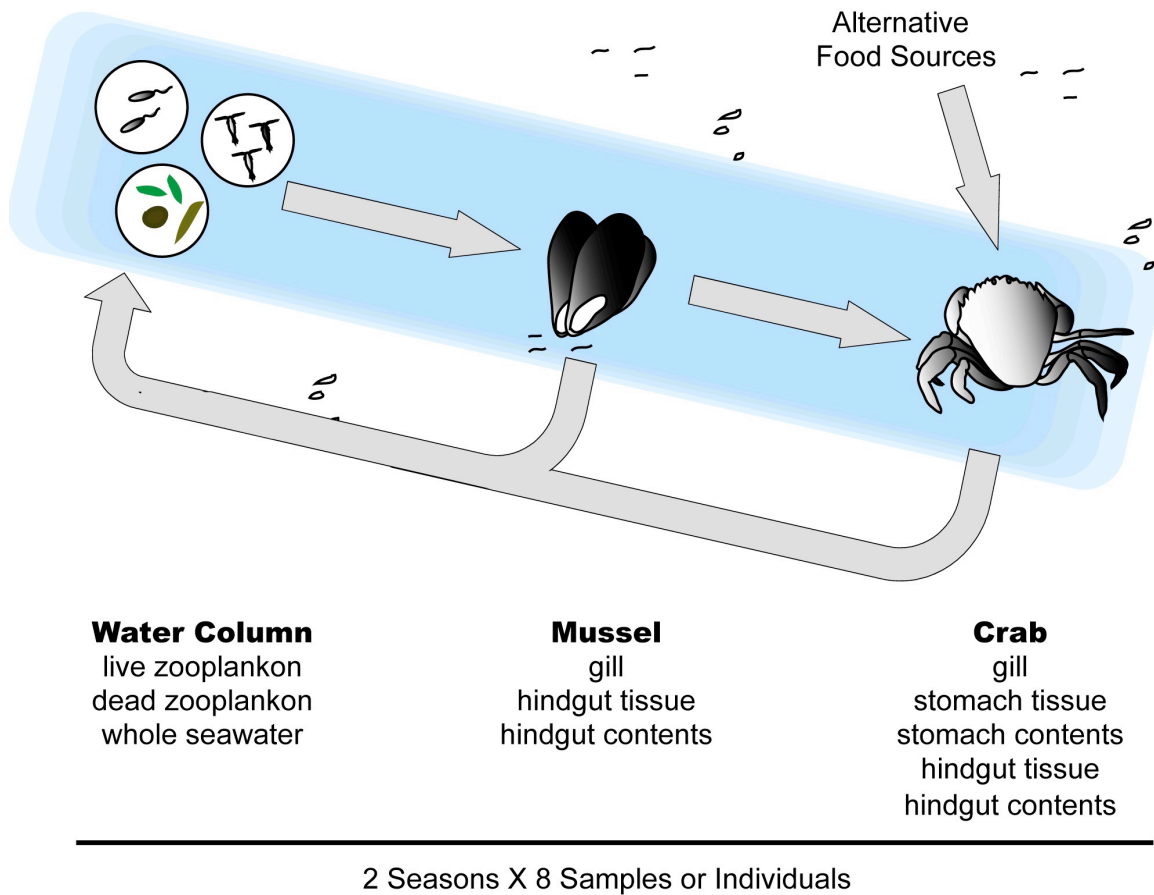


Fig. 1. Sampling strategy for comparing population structure of *Vibrionaceae* bacteria across animal hosts and water column samples in two seasons (fall and spring). Arrows indicate potential dispersal routes of bacterial populations via food chain interactions: mussels filter water (containing zooplankton) and crabs preferentially consume mussels but have alternative food sources. The blue background color indicates potential colonization via seawater. For crabs and mussels, eight individuals in each season were dissected to collect respiratory and digestive tissue, as well as stomach and gut contents. For zooplankton, an average of about 50 live or dead individuals were pooled per each of eight samples.

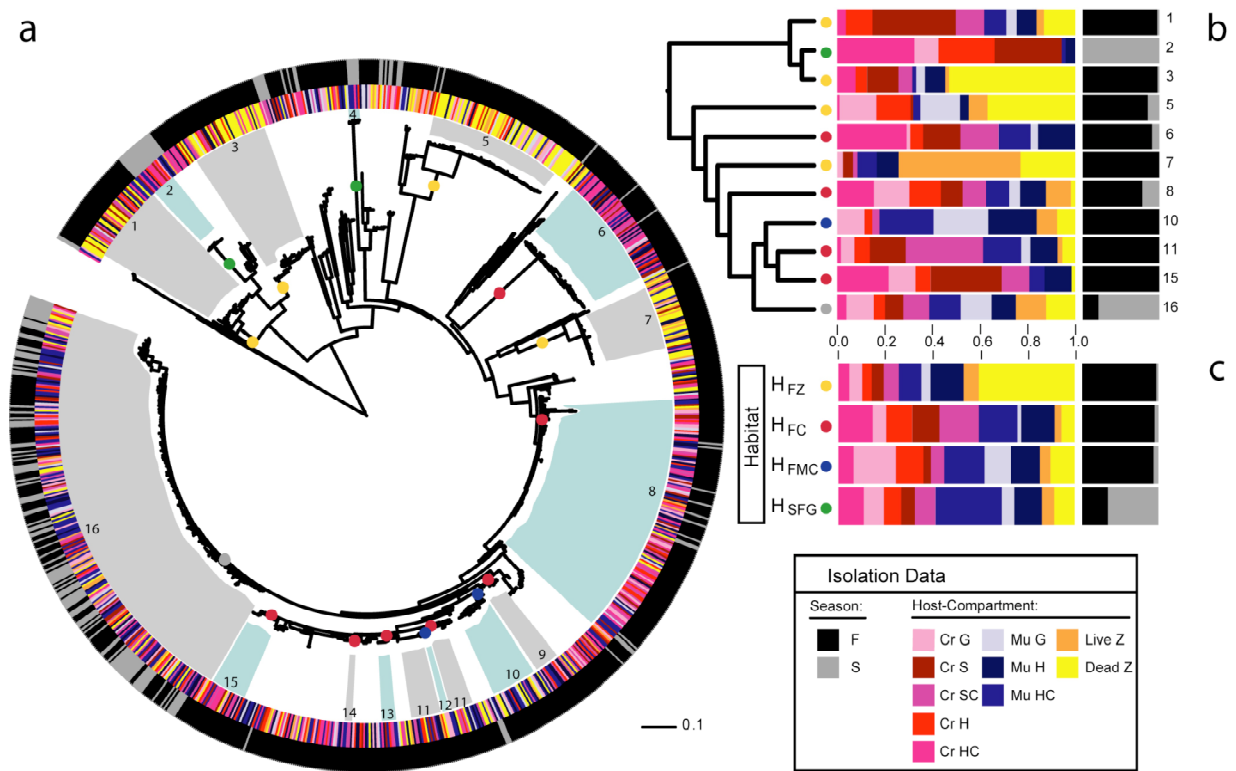


Fig. 2. Population structure predictions of *Vibrionaceae* bacteria recovered from different types of animal samples. (a) Phylogenetic tree of 1,753 strains (clones from the same sample condensed into a single sequence) denoting season and host categories from which each strain was obtained (outer ring – season; inner ring - host-compartment). Ecological populations predicted by the AdaptML algorithm (Hunt *et al.*, 2008c) are identified by alternating grey and blue shading if they pass an empirical confidence threshold of 99.9%. Population #16 was identified as a generalist (see text). (b) Tree summarizing the distribution of each population across seasons and animal samples with populations >20 members. (c) Inferred habitats from AdaptML labeled as H_{FSG} (fall-spring generalist), H_{FZ} (fall-zooplankton), H_{FC} (fall-crab) and H_{FCM} (fall-crab-mussel). Population distributions in (b) were normalized by the total number of counts in each

category to reduce the effect of uneven sampling however the habitat distributions in (c) are directly predicted from the algorithm. Identical sequence types (STs, strains identical in all 3 housekeeping genes) from the same sample, possibly representing clonal expansions, were condensed for analysis in (a) but were included in (b). Isolation data abbreviated as follows: F, fall; S, Spring; Cr, Crab; Mu, Mussel; G, Gill; S, Stomach Tissue; SC, Stomach Contents; H, Hindgut Tissue; HC, Hindgut Contents; Z, Zooplankton. Groups correspond to the following named taxa: G1, *Aliivibrio fischeri*; G2, *Aliivibrio logei*; G3, *Allivibrio sp. I3*; G4, *V. ordalii*; G5, *Enterovibrio calviensis*-like clade; G6, *V. breoganii*; G7, *Vibrio sp. F10*; G8, *V. crassostreae*; G9-G15, *V. tasmaniensis*; G16, *V. splendidus*. Scale bar is in units of nucleotide substitutions per site.

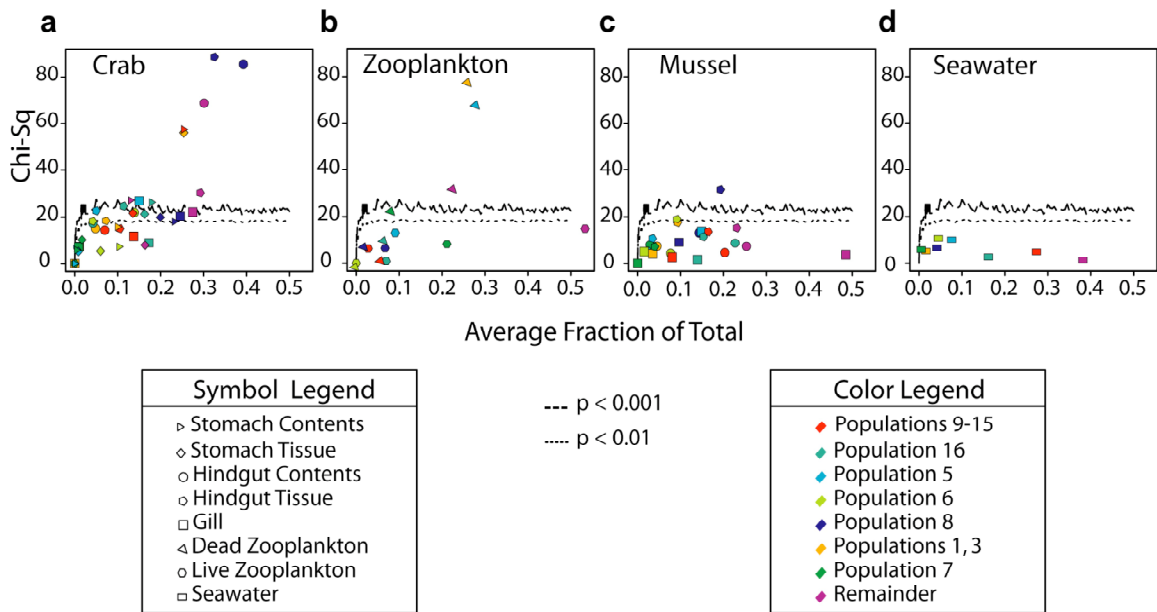


Fig. 3. Variation in the distribution of population fractions among different host specimens (individuals) or pooled samples (zooplankton and seawater). A single value, chi-squared, is used to test how each vibrio population (as a fraction of total vibrios) is distributed among tissue or tissue content samples from different individuals (zooplankton and mussel) or among pooled replicate samples (zooplankton and seawater). Chi-square values used to test for homogenous proportions are plotted against the average population fraction in (a) crabs, (b) zooplankton, (c) mussels, and (d) seawater. Symbols represent each compartment (tissue or tissue content) or sampling category (pooled sample) and are colored to identify populations. Values lying above the $p = 0.01$ (dotted line) and $p = 0.001$ (dashed line) indicate uneven population distribution among individuals. Critical values were obtained from 5000 bootstrap replicates.

SUPPORTING INFORMATION

Habitat Stability

To assess the frequency with which different numbers of habitats are predicted, 100 independent trials of the iterative habitat learning process of AdaptML were performed (see Hunt, David et al 2008 for detailed description of algorithm) (Table S1). This shows that predictions of 4 and 5 habitats occurred with nearly equal frequency while other habitat predictions were much lower. The decision to use 4 habitats instead of 5 was based on the following analysis. We compared the probability distributions of one output with four and five habitats to the remaining trials, respectively. This showed that three of the habitats were stable occurring at similar, high frequencies, while the fourth habitat (H_{SFG}) was split under the five-habitat regime (Tables S2). Because the split resulted in a fifth habitat with much lower frequency of occurrence, we chose an output matrix with four ‘habitats’.

SUPPORTING REFERENCE

Hunt DE, Lawrence AD, Gevers D, Preheim SP, Alm EJ, & Polz MF (2008) *Science* **320**, 1081-1085.

SUPPORTING FIGURES AND TABLES:

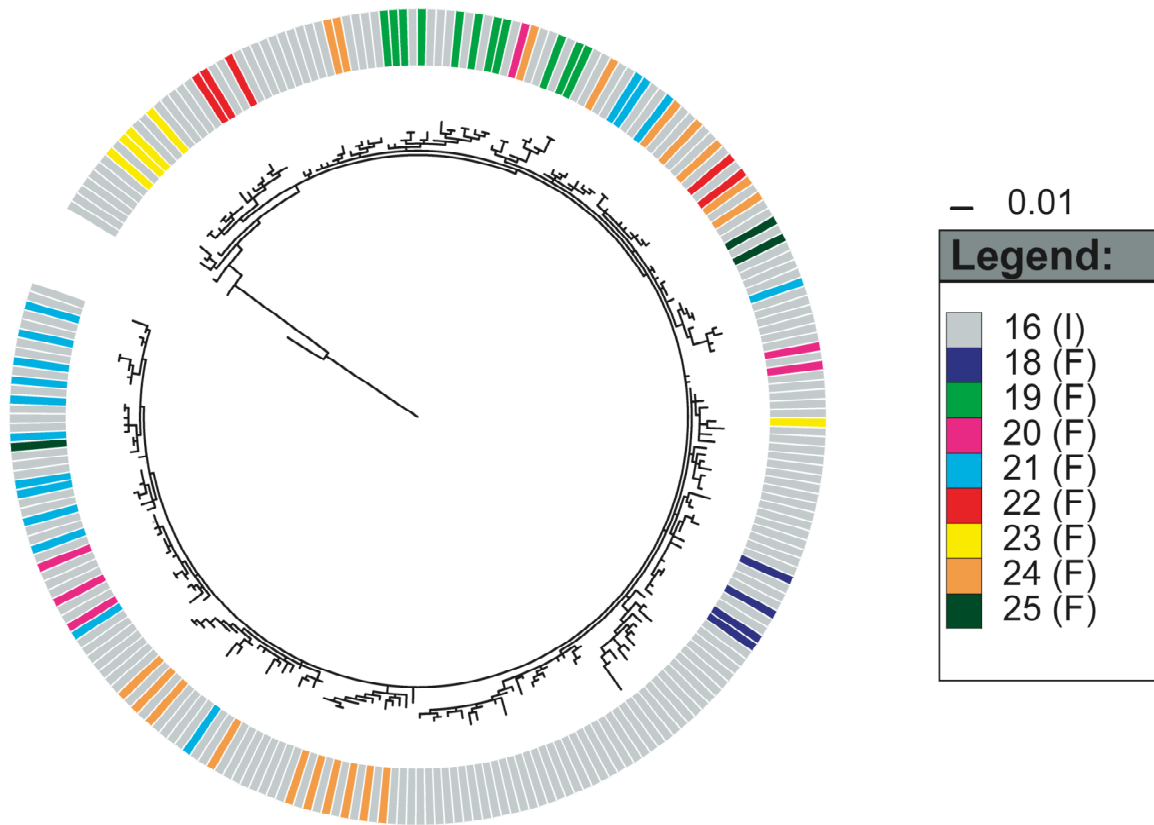


Fig S1. Comparison of phylogenetic relationships of *V. splendidus* populations #18-25 from Hunt, David et al. 2008 and *V. splendidus* population #16 from this study demonstrating that these are interleaved and indistinguishable based on maximum likelihood estimation of the hsp60 sequence. The assignment of each strain to a population is indicated by the colors on the outer circle. Population numbers in the legend correspond to populations from this study 16 (I) or in Hunt, David et al. 2008 (F). Clones were removed for analysis. Scale bar is in units of nucleotide substitutions per site.

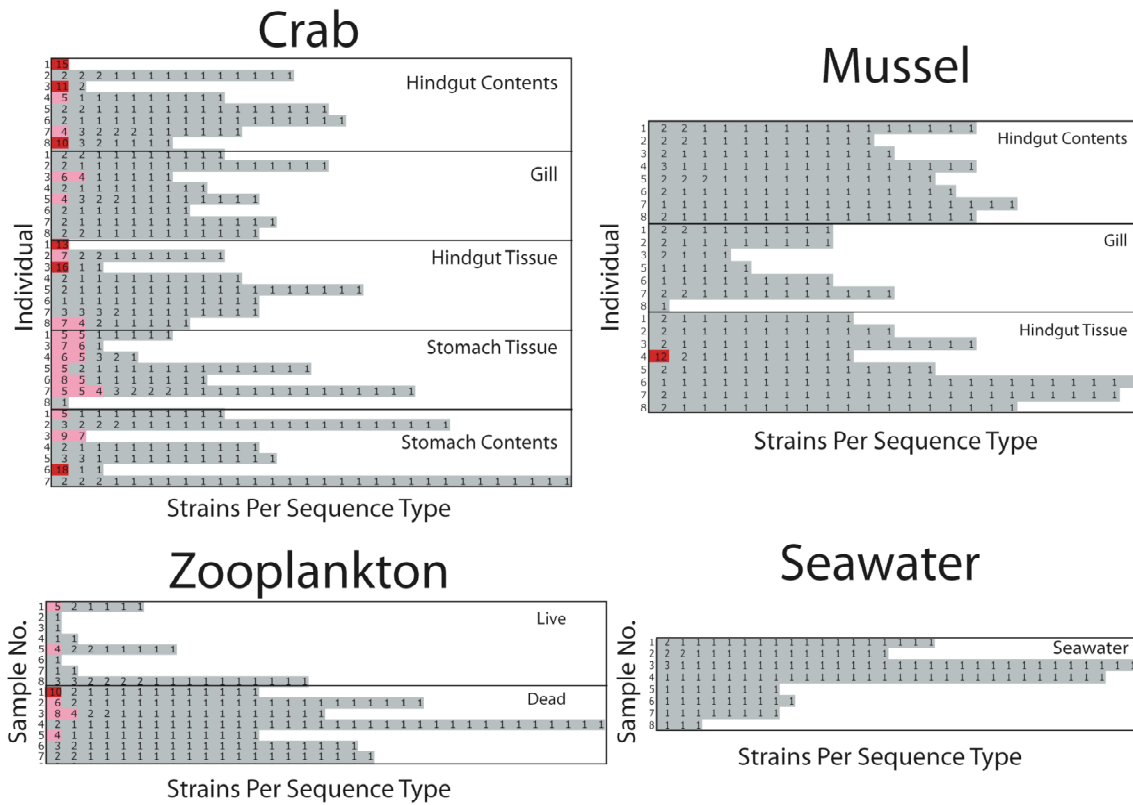


Fig. S2. Comparison of number of strains per sequence type (ST) for each sample of crab and mussel compartments, aggregate zooplankton, and seawater. Crab and mussel samples were from 8 single specimens, while the 8 zooplankton samples contained ~50 live or dead specimens, and seawater samples represented subsamples (5 μ to 5 ml) from eight 4L samples. Colors indicate isolates per ST: grey = 1-3, pink = 4-9 and red = >10.

Table S1. Reproducibility of the total number of habitats resulting from one hundred independent trials, which predicted three to six total habitats with the percent distribution across those trials as listed above.

Number of predicted habitats	Distribution across independent trials (%)
3	12
4	45
5	42
6	1

Table S2. Reproducibility of inferred habitats.

Habitat	Frequency of habitat occurrence ¹
H _{SFG}	0.707 ²
H _{FMC}	0.879
H _{FC}	1.00
H _{FZ}	1.00

¹ The frequency of habitats shown in Figure 2c using a <0.0185 average emission probability difference across 99 independent trials to identify matching habitats.

² Forty-two percent of the trials split H_{SFG} into two groups, one of which matched H_{SFG} and had a similarly high probability of occurrence, while the fifth and new habitat had a much lower probability (0.42).

Chapter Three

Resource Partitioning and Substrate Pre-Colonization Likely Influence

***Vibrio* Population Structure on Water Column Particles²**

² To be submitted with Alison Takemura, Yan Boucher, Hans Wildschutte and Martin F. Polz

ABSTRACT

Substantial physical and chemical structure lies within one liter of seemingly homogeneous seawater, inviting bacterial specialization. Several studies have demonstrated that bacterial communities attached to particulate material suspended in seawater are different from unattached bacterial communities; however, it is unknown whether specialization for chemically distinct particles commonly occurs among co-existing bacterial populations. Here we demonstrate that *Vibrio* populations can be highly specialized not only for particles of a specific type (plant- or zooplankton-derived), but also for specific substrates within one particle type. Approximately 600 isolates were obtained from particles of either plant (294 strains), or zooplankton (307 strains) origin across two seasons. Partial sequences from three housekeeping loci (*adk*, *mdh*, *hsp60*) were used for phylogenetic characterization. We used a mathematical and statistical modeling algorithm (AdaptML) to identify habitat preferences and bacterial populations of these isolates. Three of the resulting populations showed strong preferences for one of the two sampled particle types. Growth profiles on 102 carbon substrates confirmed extensive differentiation of substrate utilization capabilities between these three populations, some which could be linked to metabolism of substrates derived from brown-algae and zooplankton. Competition experiments between these populations were done *in vitro* on sterilized native particles to further investigate competitive interaction. Our results suggest that abundant bacterial populations that co-exist on the same particles seem to partition substrates more strongly than species not often found on the same particle type, providing evidence for competitive exclusion. Resource partitioning and

growth of specialist species on various substrates, along with environmental distributions of species on living substrates before death and degradation into particles can be used to explain the distribution of these species on particles in the environment. This research supports the principle that competitive exclusion is important in shaping metabolic profiles of bacteria, while suggesting that metabolism alone is insufficient to explain observed environmental distributions of bacterial populations.

INTRODUCTION

Substrates that support bacterial growth in the ocean occur either as dissolved (DOM) or particulate organic matter (POM) (Ducklow & Carlson, 1992; Simon *et al.*, 2002) and free-living and attached bacteria have been shown to be phylogenetically distinct suggesting differential specialization (Acinas *et al.*, 1999; Crump *et al.*, 1999; Delong *et al.*, 1993). However, it remains unknown to what extent bacterial populations partition particulate resources by specifically attaching to and growing on chemically and physically distinct solids. Differences between surface-attached and free-living communities may result from resistance to protozoan grazing (Matz *et al.*, 2005) and viral lysis (Filippini *et al.*, 2006) due to biofilm formation of attached cells. On the other hand, bacterial association with distinct types of particles may result from specialized attachment to and degradation of surface-associated molecules or specific interactions between attached microbial members. Particles in the euphotic zone of lakes and oceans are mostly (>90%) comprised of both living and dead bacteria, algae and zooplankton (Bishop *et al.*, 1977; Gibbs, 1974). While some chemical similarities exist across

particles of all types, those derived from macro-algae are rich in alginate and algal-specific storage components while zooplankton has a chitinous exo-skeleton. This makes these types of particles a rich and realistic experimental system with which to test the ecological mechanisms by which populations assemble on different types of particles.

Most studies have focused on delineating the total bacterial diversity within an ecosystem while relatively little is known about mechanisms of community assembly (Martiny *et al.*, 2006). Two factors may be important determinants of microbial community composition; (i) environmental conditions (selection) and (ii) migration (gene-flow). If the environmental conditions determine assembly, such that substrate specific attachment and/or growth controls the community composition, niche-based concepts of community assembly can explain the co-existence of populations (Tilman, 1982). Alternatively, migration can either be limited and generate bottlenecks (Fraser *et al.*, 2009) or be so high that it outpaces adaptation leading to more stochastic or neutral community assembly (Sloan *et al.*, 2006). It will be important to determine the relative contribution of each of these factors on the community composition to determine how predictably microbial communities assemble given certain conditions.

For particle colonization in the water column, it remains poorly understood how selection for specialization and migration are balanced. If particles are sufficiently stable and migration limited, species with large overlap in resource requirements will be less likely to co-exist because of competitive exclusion (Tilman, 1982). Over evolutionary time scales, competition can cause species to adapt alternative life-styles, e.g., to attach to different types of particles or to utilize different substrates on the same particle. Such

trade-offs are a necessary result of constraints on cellular machinery, such as cell-surface proteins and enzymatic capacity, or genome size, resulting in specialized populations with limited niche overlap. However, it is not known to what degree the water column represents a sufficiently stable environment for specialization to develop. Over 75% of particles in the water are the remains of algae and zooplankton (Bishop *et al.*, 1977; Gibbs, 1974), which, when alive, possess a suite of mechanisms to discourage bacterial colonization. When these die, they may be preferentially colonized by strains, which randomly come in contact with particles or are close enough to chemotax towards them (Seymour *et al.*, 2009). Additionally, bacteria resistant to defense mechanisms may pre-colonize the living organism and persist. Given these possibilities, it remains an open question how specifically and predictably populations assemble on different types of particles.

Previously, it was shown that *Vibrios* co-occurring in the water column can be divided into populations, which are either relatively more abundant on particles of specific size classes or free-living in the water column (Hunt *et al.*, 2008a). Although this is suggestive of them being attached to particles of different composition, this has yet to be demonstrated. Here, we investigate the distribution of *Vibrios* in the environment on different, identifiable types of particles, which dominate the POM in coastal seawater: zooplankton and plant-derived particles. Further, we use growth assays and competition experiments on both particle types to investigate how specialization and migration might explain the observed environmental distribution. Our results suggest a high degree of specialization for particles and particle-specific substrates of co-existing vibrio

populations, supporting trade-off based niche theory of community assembly. However, based on results from direct competition assays, historical effects (i.e. colonization of living organism prior to death) might be an important factor needed to explain observed environmental distributions.

RESULTS

Environmental Distribution

Probing the environment for bacterial population structure can provide information about specialization for different types of POM. We compared the distribution of over 600 *Vibrio* strains from two categories of microscopic POM: plant-derived (long and globular) and zooplankton (living and dead) particles. Particles were picked by eye under a dissecting microscope from the >64 μM size fraction collected from 100 L of seawater. Three housekeeping gene sequences and sampling categories served as inputs for a model to predict bacterial population structure among the different particle types [i.e., identify groups of closely related strains with characteristic distributions ('habitats') among the sample categories] (Hunt *et al.*, 2008a). In the following, we limit the analysis to populations for which more than 10 isolates were recovered.

Our analysis suggests that the particle-attached *Vibrio* strains can be grouped into eight populations (Fig. 1A), whose distributions among the sample categories (Fig. 1B) are summarized by four characteristic 'habitats' predicted by the model (Fig. 1C). One of these habitats, H_{SFG}, is characterized by nearly equal relative abundances on particles of

all types across both seasons. Remaining predicted habitats were limited to the fall. Two habitats, H_{FP} and H_{FZ} , had distributions that were skewed in favor of either plant material or zooplankton, respectively. The final habitat H_{FG} was an intermediate between the two specialized habitats, with a slight preference for zooplankton over plant-derived particles.

Mapping the distribution among particles onto predicted populations suggests a range of specialization. All but one population was detected on all particle types, suggesting either relatively high migration between particles or weak specialization. Two co-occurring *Vibrio* populations (*V. splendidus* and *V. tasmaniensis*) had distributions across particle types consistent with an ecological generalist; this is predicted from the model (H_{SFG} and H_{FG} ; Fig. 1C) and shown in their normalized distributions (Fig. 1B). Three populations demonstrate a slight preference for dead zooplankton (*Enterovibrio calviensis*-like, *Aliivibrio fischeri* and *Allivibrio* sp. I3; Fig. 1B) while both the model predictions and normalized distributions suggests a high degree of specificity for three additional populations for plant-derived (*V. breoganii* and *V. crassostreae*) and zooplankton (*Vibrio* sp. F10) particles (Fig. 1B). *V. breoganii* was the only significant population completely excluded from both living and dead zooplankton particles. All populations were present on elongated or globular plant-derived particles, if even at a low relative frequency. Only one population was relatively more abundant in the spring (*V. splendidus*). This population seems to gain a significant advantage over other *Vibrio* populations at lower temperatures in the spring, allowing it to dominate across all particle types. In the fall, its distribution remains nearly equal across all particle types, but it is not the most relatively abundant population.

Colonization of the living organism by specific *Vibrio* populations prior to death might account for some of the observed environmental distributions. While vibrios were isolated from both living and dead zooplankton in this study, none of the particles consisted of living plant material. We therefore tested whether some of the populations occur on locally abundant brown algae (rockweed, *Fucus vesiculosus*), which are located at the same site and a potential source of plant-derived particles. Brown algae fronds were homogenized and plated on *Vibrio* selective media. While no vibrios were detected on rockweed samples collected in the spring, many populations isolated from these algae in the fall (Fig. 2). It is unknown what percentage of plant-derived particles stems from the remains of these brown macro-algae. However, it is suggestive that the three populations that dominate on brown algae fronds (*V. splendidus*, *V. crassostreae* and *V. breoganii*) are also the three most abundantly populations on globular plant-derived particles. Further investigation is warranted to determine potential other sources of plant-derived particles.

Metabolic Differentiation of Populations

Population-level variation in metabolic capabilities may be one mechanism creating differential patterns of occurrence on particles, as predicted by niche-based theories of microbial assembly. We therefore tested growth on substrates, which are differentially distributed among particle types: chitin for zooplankton, and alginate and cellulose for plants. Additionally, fine-scale metabolic differentiation was assayed for populations with the same particle preferences. Four strains from each of the three

populations with the most strongest particle preferences (*V. breoganii*, *V. crassostreae*, and *Vibrio sp. F10*) were used for growth assays on 95 standard carbon sources on Biolog GN2 plates and on chitin, alginate and various cellulose derivatives.

The three particle specialist populations have overall very different substrate utilization profiles, suggesting metabolic resource partitioning. Average Biolog correlation coefficients within populations are significantly higher than between populations (within, 0.853; between, 0.486; $p > 0.001$), as would be expected if most traits were shared within populations (Table 1). Between the plant-particle specialist populations, *V. breoganii* and *V. crassostreae*, the latter was able to metabolize more amino and carboxylic acids (D-Alanine, L-Asparagine, L-Aspartic Acid, L-Histidine, L-Proline, L-Serine, Succinic Acid, Bromosuccinic Acid, alpha-Keto Glutaric Acid, Acetic Acid, Cis-Aconitic Acid and Citric Acid). Carbohydrates also differentiated the populations with *V. breoganii* being able to utilize Dextrin and Gentiobiose while *V. crassostreae* could metabolize D-Trehalose, D-Mannose and D-Galactose (Table 1). In general, *Vibrio sp. F10* did not possess any population-specific metabolic capabilities detectable by Biolog assays. Instead, *Vibrio sp. F10* appears to have lost the ability to metabolize substrates that other populations possessed.

Of the particle-specific substrates, alginate degradation differentiated both plant specialist from the zooplankton specialist populations, but ability to degrade chitin did not correspond to zooplankton preference. None of the isolates showed growth on any of the cellulose substrates tested. Both populations with plant-derived particle preferences could metabolize alginate, but *Vibrio sp. F10*, which has a zooplankton preference, could

not. Only *V. crassostreae* was able to grow on chitin, even though this property is widely conserved across the vibrios (Hunt *et al.*, 2008b) and *Vibrio* sp. F10 appears more specific for zooplankton than *V. crassostreae* with 30/37 and 12/77 isolates obtained from zooplankton particles out of the total number of isolates, respectively. On the other hand, *V. breoganii* was never isolated from zooplankton (0/65 total isolates) so that the absence of chitin metabolism appears less surprising. Thus, chitin degradation does not correspond to presence on zooplankton, while alginate degradation is a good predictor of presence on plant-derived particles.

In vitro Competition for Particle Attachment

In the environment, the relative abundance of populations on particles may be affected by two factors: (i) preferential attachment and growth and (ii) migration or dispersal. Migration or dispersal could create elevated cell counts despite low metabolic activity in the present environment as recently demonstrated for thermophiles in Arctic water (Hubert *et al.*, 2009). Additionally, historical effects, such as pre-colonization of the particle, could compensate for possible growth advantages of one population over another. However, if the growth advantage of one population over the other is very strong, even very high migration may not be enough to overcome the difference in fitness. We investigated effects of relative fitness by competing pairs of strains representing different populations for mixtures of sterilized zooplankton and plant-derived particles. We also simulated the effects of high migration by altering the inoculation ratios of the populations to determine how these changes affect the relative

population distribution on particles of each type. For each trial, two strains, each from a population with specialized particle preferences predicted from environmental distributions, were competed on mixtures of ten and fifteen sterilized plant-derived and zooplankton particles in seawater without nutrient amendment. DNA was extracted from particles after up to four-week incubations at room temperature. The relative proportions of each population on the particles of each type were analyzed by quantitative PCR and measured by constant denaturing capillary electrophoresis (Thompson *et al.*, 2004). Competition of populations with the same, or opposite particle preferences, along with either replicates of the same strain pair, or different strains within the same population provide information of the reproducibility of the four-week outcome. The dynamics of the system were also determined by looking at additional one- and/or two- week time points, providing information about how the system is changing over the course of the experiment.

In a pilot experiment, the one *Vibrio sp. F10* and *V. breoganii* strain, representing zooplankton and plant specialist populations, were competed in triplicate incubations to estimate reproducibility of the competition (Fig. 3). On plant-particles, although there was no statistical difference at week one, from week two onwards, *V. breoganii* was significantly more abundant than *Vibrio sp. F10* ($p < 0.001$ for both 2- and 4 week time points). On zooplankton particles, the zooplankton specialist (*Vibrio sp. F10*) was relatively more abundant than *V. breoganii* only after incubation for four weeks ($p < 0.01$). In seawater alone (no particles), the abundance of *Vibrio sp. F10* is significantly greater than *V. breoganii* ($p < 0.01$) after one week; however, this difference is not significant in

the two- and four- week time points. This suggests that replicate incubations of the same strain pair display similar relative abundances after 4 weeks and that the relative abundance of each population is consistent with the observed abundance of that population on the particles found in the environment.

Since strain- specific variations in genome content has been shown to be extensive among closely related bacteria (Thompson *et al.*, 2005; Welch *et al.*, 2002) and may lead to different fitness values, we extended the analysis to include different strains of the same population (biological replicates). Three to four different strains pair combinations were used in the rest of the *in vitro* competition assays to better estimate population-wide competitive abilities.

Zooplankton vs. plant particle specialist (*Vibrio sp.* F10 vs. *V. breoganii*): When different strains of *V. breoganii* and *Vibrio sp.* F10 were used as biological replicates for the competition assay, the average distribution of each population across particle types supports the environmental distribution; however, the standard deviations are very large such that they are not significantly different (Fig. 4D-F). *V. breoganii* is significantly more abundant on plant-derived particles than *Vibrio sp.* F10 at two weeks ($p < 0.001$); however, at the four week time point, *V. breoganii* is not significantly more abundant despite having similar average values as the two week time point. On zooplankton particles, there is no statistically significant difference in relative abundances of either population, even after four weeks. Population wide average values suggest that *Vibrio sp.* F10 is more relatively abundant on zooplankton particles than *V. breoganii* on average,

but the outcome for each strain pair varies such that this average is not significant. No systematic bias was seen in strain pair combinations (Fig. 5). While the environmental distribution was supported by the pilot experiment with replicates of one pair of strains, using biological replicates did not provide statistical support for the environmental distributions.

When *Vibrio sp. F10* was inoculated into the assay at 100-fold higher relative abundance over *V. breoganii*, it remained significantly more abundant on both particle types after four- week incubations (Fig. 4A-C). Similarly, *V. breoganii* was also the most relatively abundant on both particle types when it was inoculated at 100-fold higher relative abundance over *Vibrio sp. F10* (Fig. 4G-I). Neither population could displace the other regardless of initial density, suggesting the competitive advantage that one population has over the other on either particle type is small, if any. Because the competitive advantage of each population type is very low, it might take more than one month for the populations to reach steady state if the distribution is primarily determined by the environmental conditions (and competitive advantage of one over the other) on the particles. Especially for zooplankton, which are degraded three days at 20-22 °C in laboratory experiments with seawater (Tang *et al.*, 2006), the relative growth dynamics of only two vibrio populations on these particles in a simple system is not sufficient to explain their distribution.

Zooplankton vs. plant particle specialist (*Vibrio sp. F10* vs. *V. crassostreae*):
Despite an observed 13-fold higher abundance of *V. crassostreae* over *Vibrio sp. F10* on

plant-derived particles in the water column, the latter was, on average, more abundant than *V. crassostreae* on plant particles after four weeks, although not significantly so (Fig. 6D). *V. crassostreae* might also be expected to dominate on dead zooplankton, since it can degrade chitin while *Vibrio sp. F10* can not. Yet on average, *Vibrio sp. F10* are also more abundant than *V. crassostreae* on zooplankton particles, although not significantly so (Fig. 6E). *Vibrio sp. F10* did not possess any assayed metabolic capability that *V. crassostreae* did not (as determined by growth on Biolog GN2 plates, chitin and alginate), making the nature of the dominance difficult to assess.

Strain- pair specific competition differences create high variability on zooplankton particles, as demonstrated by changing the initial density of both species. On plant-derived particles, starting with 100-fold higher initial density of one population resulted in the more abundant population being statistically significantly more abundant after four weeks (Fig. 6A and G). Interestingly, the less abundant population increased from ~1% to ~25% of the total population in both cases. On zooplankton particles, changing the inoculation ratios changed the average population relative abundances, although again, these differences were not significant (Fig. 6B and H). Strain to strain variation accounts for the large standard deviation in average population ratios at the four-week time points. In one strain pair, *Vibrio sp. F10* was always dominant at four weeks, regardless of initial density of either population (Fig. 7, #6). Yet in another strain pair, *V. crassostreae* is always dominant after four-weeks (Fig. 7, #8). In the last biological replicate, when *V. crassostreae* is added at 100-fold higher initial density, it is more abundant at four weeks, which is different from when inoculated at equal or lower

abundances than *Vibrio sp. F10* (Fig. 7, #7). This high strain-to-strain variability makes it difficult to draw any conclusions about population- wide trends in competition for these particles.

Plant vs. plant specialist (*V. crassostreae* vs. *V. breoganii*): Populations with the same predicted particle preferences were tested on both zooplankton and plant-derived particle types. As predicted from environmental data, *V. crassostreae* was significantly more abundant than *V. breoganii* on zooplankton particles already after two weeks (Fig. 8E). Additionally, adding 10- or 100- fold more *V. breoganii* to the initial inoculum did not alter the final outcome after four weeks (Fig. 8B). Since *V. crassostreae* can metabolize chitin and many amino acids that *V. breoganii* cannot, these results are consistent with near complete loss of these functions in *V. breoganii*.

On plant-derived particles, both *V. crassostreae* and *V. breoganii* are present at close to equal relative abundances when inoculated at similar starting ratio (Fig. 8D). *V. crassostreae* is significantly more abundant than *V. breoganii* on these particles when inoculated at 10- and 100-fold higher initial density (Fig. 8G). When *V. breoganii* is inoculated at 10- or 100-fold more, these species have nearly equal ratios on plant-derived particles after just two weeks (Fig. 8A).

These results can be explained in a non-equilibrium system by a loss of function in *V. breoganii*. These two populations might co-exist on these particles by binding and/or metabolizing one of two equally abundant substrates (e.g. amino acids and alginate, for illustrative purposes). At equilibrium, each population is expected to be present at a concentration relative to the abundance of its preferred substrate, since each

would eventually out-compete the other for these substrates. In a non-equilibrium state, where the amount of another limiting resource, such as space on the particle, is diminished before equilibrium on these substrates is established, the rate at which populations can out-compete the other for its preferred substrate may cause different outcomes. If *V. breoganii* lost the ability to bind and/or metabolize its non-preferred substrate (many amino acids) but *V. crassostreae* did not, then the rate at which *V. crassostreae* binds and/or grows on its non-preferred substrate (alginate) will be much higher than the rate at which *V. breoganii* binds its non-preferred substrate (amino acids). When changing initial densities of both populations, the difference in rates could create a pattern similar to what was observed.

Another possibility is cross-feeding of *V. breoganii* by *V. crassostreae*, since *V. breoganii* can not achieve higher relative abundance on plant particles than *V. crassostreae*. This may parallel the co-existence of polymorphic strains of *E. coli* in a chemostat through differential secretion and uptake of glucose, acetate and glycerol substrates (Rosenzweig *et al.*, 1994). Cross-feeding is supported by the environmental data (Table 2). In each of the eight replicate particle samples collected in the fall, *V. breoganii* is only present if *V. crassostreae* is also relatively abundant. *V. crassostreae* reaches relatively high abundances in four samples where *V. breoganii* is not found or only present at very low abundances. In the case where *V. breoganii* has lost the ability to metabolize its non-preferred substrate, one might expect to see *V. breoganii* abundant on particles alone, since the possibility exists that other organisms aside from *V. crassostreae* might also compete for the non-preferred substrate, displacing *V.*

crassostreae while *V. breoganii* is still present. Since a mechanism of cross-feeding is speculative, further experiments are needed establish the mechanism of their interaction.

Strain-Specific Antagonistic Interactions

Antagonistic interactions could alter the competitive ability of each strain pair tested on particles of each type via interference competition (Long & Azam, 2001). We tested pair-wise antagonistic interactions for each of the biological replicates used in the *in vitro* competition assays. Briefly, interaction between strains can be determined if the presence of one strain inhibits the growth of the other on rich media (blood-heart infusion agar plates). In this way, many strain-specific interactions were evident (Table 4), two of which involved strain-pairs used in competition assays. Neither of the potential producer strain dominated in competition assays regardless of initial density, as might be expected of interactions involving this type of antagonistic interaction. Interestingly, all of the *Vibrio* sp. F10 strains inhibited growth of at least one *V. breoganii*. Whether these interactions translate into similar antagonistic interactions of particles of each type is unknown.

DISCUSSION

Understanding the factors controlling population distributions and abundances is a primary goal of ecological research. Two factors may weigh prominently in determining which populations are abundant in a given environment; (i) influx of strains, either historically or presently, into the environment and (ii) competition for space or resources.

We first showed the environmental distribution of vibrio populations across particles of different types in the water column by conservatively predicting their preferred habitats based on preferential isolation from different particle types. Picking populations with the most uneven environmental distributions, we then tested the competitive ability of populations with the same or opposite predicted habitat preferences on particles from which they were originally isolated. We were able to probe further into the nature of the competition by altering the inoculation ratios. Three outcomes were clear from the data. Altering the initial density of populations changed relative population abundances after four- week incubations. Assuming a non-equilibrium state, these results are consistent with one population having only a low competitive advantage of one population over the other, which can be overcome by the addition of a higher density of cells. Second, strain-pair interactions were vastly different between biological replicates. Changing the inoculation ratio did not change the results of competition after four weeks in two of three cases, but the dominant population was different in each case. Finally, one population could not out-compete the other regardless of initial starting density. Overall, this approach is unique in its scope and has the potential to test ecological theories under somewhat realistic conditions using microorganisms.

Model results from environmental sampling data predicted four habitat types with respect to particle preference across eight different populations within the *Vibrionaceae*. One habitat preference suggested a generalist population with respect to particle type (H_{SFG}); however, this population was relatively more abundant in the spring, thus being seasonal differentiated with respect to all other *Vibrio* populations. Four populations were

predicted to be relatively more abundant on zooplankton particles, with varying degrees of association with plant-derived particles. Of these, *Vibrio* sp. F10 had the highest relative association with living zooplankton, with other zooplankton populations possibly having a more saprophytic lifestyle. The plant-derived habitat preference was only predicted for two populations, making zooplankton the generally preferred habitat for vibrios. This is consistent with the conservation of the chitin degradation pathway across the *Vibrionaceae* (Hunt *et al.*, 2008b). Interestingly, only one of the three populations chosen for further investigation in this study was able to degrade chitin, and chitin degradation did not correspond to presence on zooplankton particles.

In this study, *Vibrio* species found in the environment on the same particle type show evidence of partitioning particle-specific substrates. This was most convincingly demonstrated for *V. breoganii* and *V. crassostreae*, which are both found on plant-derived particles in the environment (Fig. 1). These species have widely different growth profiles on chitin and algal-derived substrates. Growth is statistically significantly different on 24 carbon substrates used as sole carbon sources, many of which are directly or indirectly related to metabolism of algal-derived substrates. Metabolism of gentiobiose, found exclusively in *V. breoganii*, might be a side-effect of metabolizing laminarin, an abundant storage molecule in brown algae (Powell & Meeuse, 1964). Dextrin, which is formed from the hydrolysis of starch, is preferentially degraded by *V. crassostreae* and along with many amino acids. Their inability to degrade cellulose, another abundant polymer found on plant derived particles, suggests that rather than these population being plant-particle specialists, their preferred habitat may be the subset of

plant-derived particles originating from macro-algae. This is supported by the high relative abundance of these populations on brown algae fronds.

Additional evidence for resource partitioning, specifically cross-feeding of one population by the other, was also seen in both *in vitro* experiments and from environmental sampling. On zooplankton particles, *V. crassostreae* outcompetes *V. breoganii* even if it *V. breoganii* is inoculated at 100-fold higher initial concentrations. On plant-derived particles, increasing the initial density of *V. breoganii* 100-fold results in the same relative abundance of both populations after four weeks, whereas increasing the initial density of *V. crassostreae* 100-fold allows it to outcompete *V. breoganii*. These results are consistent with a scenario of co-existence where competitive exclusion has resulted in a strong degree of resource partitioning, either by exclusive metabolism of different substrates or through some cross-feeding mechanism. Evidence from sampling data suggest a cross-feeding scenario, since *V. breoganii* is only at high relative abundances when *V. crassostreae* is also abundant, but *V. crassostreae* is abundant when *V. breoganii* is not. Stable co-existence of these strains may mechanistically mirror the experimental evolution experiments with *E. coli* where initially isogenic strains developed stable polymorphisms, which enabled them to co-exist via differential secretion and uptake of acetate and glycerol (Rosenzweig *et al.*, 1994). Overall, the *Vibrio* population distributions in the environment and interactions in the laboratory suggest that competition has had a strong influence on their metabolic capabilities.

The interactions between *V. crassostreae* and *Vibrio* sp. *F10* were dependent on the specific strain-pair combination used in the analysis, making it difficult to draw

strong population-wide conclusions from the data. *V. crassostreae* degrades chitin, alginate and many amino acids that *Vibrio sp. F10* does not, but it is unknown which (if any) substrates are specifically metabolized by *Vibrio sp. F10*. In spite of the ability of *V. crassostreae* to degrade many substrates that *Vibrio sp. F10* could not, *Vibrio sp. F10* had on average a higher relative abundance over *V. crassostreae* on both plant-derived and zooplankton particles *in vitro*. These slightly higher average values were not significant because of the large standard deviations from specific strain pair interactions, especially on zooplankton particles. Plant-derived particles did not display the same degree of strain specific outcomes. Some strains of *Vibrio sp. F10* may specifically antagonize strains of *V. crassostreae* on chitin or vice versa; however, this interaction was not apparent on agar plates. Antibacterial agents have been shown to bind to alginate, reducing their effectiveness (Nichols *et al.*, 1988). Thus, if antibacterial agents, such as bacteriocins, secreted by one strain control interaction with others on chitin particles, the effects might be dampened on brown algae particles with high alginate content. It is unclear what role, if any, these strain-specific interactions play in the ultimate distribution of these populations on environmental particles.

Migration could influence the environmental distribution of bacterial populations, especially in the marine environment where the water offers a high degree of connectivity. Persistence on particles could be based on the relative abundance of the population in the water column when the particle surface was created, e.g., in the case of zooplankton carapace formation. Additionally, colonization of living host organisms, possibly even at low levels, could provide a significant advantage of one population over

another upon death and degradation of the host into particulate organic material. Neither population predominantly isolated from freshly collected plant particles could statistically out-compete the zooplankton-particle specialist population in *in vitro* competition experiments, unless initially inoculated at 100-fold higher relative abundances. Thus, the presence of both plant-particle specialists on living brown algae fronds, before death and degradation into particulate organic carbon, could be necessary to explain the relative abundance of these populations on plant-derived particles.

Since natural particles, characterized only by shape and color, were used in these experiments, particle heterogeneity could have account for either the large standard deviations across biological replicates, or weak relative fitness values between populations. Plant-particle specialists showed a strong degree of resource partitioning for brown algae substrates. Thus, these populations could be specialized for only the fraction of plant particles that were derived from brown algae. Only ten particles were chosen to use in competition experiments; by chance, some incubations may have had a larger fraction of plant particles derived from algae than particles derived from other plants, such as marsh grass or leaf litter. Additionally, synergistic effects with other populations, such as those provided by a mechanism of cross-feeding, may be required for the true growth advantage of the specialist population to become apparent in competition assays. Thus, while these competition assays did not show a remarkable difference in fitness values between specialist populations on their preferred particle type, an improved experimental design could provide different results.

Experimental results suggest resource partitioning of particle substrates and supports a role for pre-colonizing living host organisms before degradation into particle organic matter in explaining population abundances on freshly collected materials. Other organisms or environmental factors, such as the presence of non-*Vibrio* organisms, predation pressure by protozoan or viruses and changes in environmental conditions will certainly affect the observed environmental distribution. More environmental sampling is needed to understand how reproducible these associations are, possibly identifying other factors that correspond with relative abundance. However, here we demonstrate that both resource partitioning and migration are likely important underlying factors determining the ultimate environmental distribution.

MATERIALS AND METHODS

Water Column Particle and Brown Algae Sampling:

Water column particles and rockweed plants were collected from Plum Island Sound Estuary, Ipswich, MA in the spring (04/23-05/03) and fall (09/24-10/04) of 2007 with environmental conditions listed in Table 4. Eight replicate brown algae (*Fucus vesiculosus*) samples were picked from the shore attached to hard substrate (mussel bed, rocks, etc.), washed with sterile seawater and placed into sterile whirlpack until processing. Particles were collected by filtering approximately 100 L of seawater through a 64 μm mesh net. Eight replicate 100 L samples were collected in each season, two samples per day. Samples were washed with sterile seawater, transferred by sterile seawater water wash into a 50 ml conical tube and placed in a cooler until processing. In

the lab, living and dead zooplankton were differentiated by eye under a dissecting microscope based on movement. Plant-derived particles were also picked under a dissecting microscope by eye based on green to brown color and elongated or globular shape. 10-140 individual zooplankton or plant-derived particles were picked from each 100 L sample (not exhaustive) into 4 ml sterile seawater in a sterile tissue grinder and the large particles were broken up with extensive grinding. One brown algae frond (approximately 1/2 inch) was removed with a sterile razor and placed into 4 ml sterile seawater in a sterile tissue grinder and broken up with extensive grinding. Each homogenized sample was serially diluted (10-fold to 10,000-fold) in sterile seawater prior to plating on *Vibrio* selective media.

Bacterial Isolation and Gene Sequencing

Dilutions of all samples were plated on *Vibrio* selective marine TCBS media (thiosulfate/citrate/bile salts/sucrose, BD Difco TCBS with 1% NaCl added) for bacterial strain isolation. Bacterial colonies were picked from the most dilute samples whenever possible. Up to 20 colonies were picked from each plant or zooplankton type from each of the eight samples (brown algae, living zooplankton, dead zooplankton, oblong green/brown particle, globular brown particle), resulting in up to 160 bacterial isolates per category per season. Bacterial strains were serially re-streaked on 1% tryptic soy broth (TSB) plates (BD Bacto with 2% NaCl added), marine TCBS plates and 1% TSB plates again to purity. For gene sequencing, bacterial strains were grown in liquid culture for 2-3 days in 1% TSB at RT with shaking; 10 µl of culture was treated with 10 µl of

Lyse-N-Go (Thermo Fisher Scientific, Rockford, IL) to prepare DNA template and cycled according to manufactures recommendations. 16S primers 27f and 1492r were used to amplify the small subunit rDNA sequence (Lane, 1991). Primers targeting the *adk* (forward: 5'- GTA TTC CAC AAA TYT CTA CTG G -3' and reverse: 5'- GCT TCT TTA CCG TAG TA -3' (Hunt *et al.*, 2008a)) *hsp60* (forward: 5' GAA TTC GAI III GCI GGI GAY GGI ACI ACI AC 3' and reverse 5' CGC GGG ATC CYK IYK ITC ICC RAA ICC IGG IGC YTT 3'(Goh *et al.*, 1996)) and *mdh* (forward primer: 5'- GAT CTG AGY CAT ATC CCW AC -3' and reverse primer: 5'- GCT TCW ACM ACY TCR GTA CCC G -3') were used to amplify and sequence part of the coding region for each gene. Additional primers were designed if the initial amplification resulted in no product (*adk*: forward 5'- GCW CCD GGY GCR GGT AAA G-3' and reverse: 5'- TAG TRC CRT CRA AYT THA GGT -3'; *mdh*: forward 5'- GAY CTD AGY CAY ATC CCW AC -3'). All of the housekeeping genes were amplified using the following PCR conditions: 95°C for 3 min.; 30 cycles of 95°C for 30 sec, 37-45°C for 30 sec and 72°C for 1 min.; 72°C for 5 minutes (annealing temp for *hsp60*: 37°C, *adk*: 40°C; *mdh*: 45°C). All genes were sequenced in one direction only. Sequencing was performed at the Bay Paul Center at the Marine Biological Laboratories in Woods Hole. Automatic base calls were trimmed and manually curated using Sequencher (Gene Codes Corp., Ann Arbor, MI) and aligned using Clustalw (Jeanmougin *et al.*, 1998) with visualization and further manual curation using MacClade (Sinauer Associates, Sunderland, MA). Sequences have been submitted to GenBank under the following accession nos.; *hsp60*: GQ988782-GQ988835 (zooplankton- spring), GQ989189-GQ989441 (zooplankton- fall), GU377809 -

GU378102 (plant-derived); *adk*: GQ990535- GQ990588 (zooplankton-spring), GQ990942- GQ991194 (zooplankton-fall), GU377515 - GU377808 (plant-derived); *mdh*: GQ992288- GQ992341 (zooplankton-spring), GQ992695- GQ992947 (zooplankton-fall), GU378103 - GU378396 (plant-derived).

Phylogenetic Analysis

Gene sequences were concatenated for phylogenetic inference by maximum likelihood using PhyML (Guindon & Gascuel, 2003). Strains identical at all loci and isolated from the same sample and particle type were included only once in the phylogeny to prevent our algorithm from identifying clonal expansions as likely habitats. The GTR substitution model and 4 rate categories (all other parameters were estimated from the data) were used for phylogenetic analysis.

Population Predictions

AdaptML was used to identify population structure (Hunt *et al.*, 2008a) using the strain phylogeny and environmental annotations (season, particle type) as input for the model. AdaptML uses a Hidden Markov Model to predict (i) the phylogenetic bounds of ecologically distinct populations and (ii) their habitat composition (distribution among environmental categories). The model threshold value was set at 0.025, *E. coli* was used as the out-group and the analysis was rerun 100 times with the same phylogenetic tree to verify the stability of the predictions (Table 5). The iTOL tool was used to visualize the model output, habitat predictions and environmental categories (Letunic & Bork, 2007).

Only populations identified by AdaptML and validated by an empirical significance test at a significance level of $p < 0.05$ (Hunt *et al.*, 2008a) were selected for further analysis, with the exception of one large group of closely related strains, which was manually assigned as a generalist population (#9 in Fig. 1). All isolates in population #9 were assigned the H_{SFG} , which predicts relative equal distribution across all particle types and seasons. Thus, even when reducing the stringency of the empirical significance test to only $p < 0.05$, few of these isolates were retained in a significant population. Instead, all isolates with the H_{SFG} habitat prediction from this group were retained as population #9. Population distributions were normalized to the average total number of isolates obtained from each sampling category in each season (75.125) to control for difference in sampling effort across both categories and seasons. Although clones from the same particle type and season were removed AdaptML analysis, all clones were included in population distributions. Distributions for populations with more than 10 members are reported.

In vitro Competition Assay

Four strains from each of the three populations with the most polarized habitat preferences (*V. breoganii*, *Vibrio* sp. F10 and *V. crassostreae*) were used in a competition assay for zooplankton and plant-derived particles *in vitro* (Table 6). The following strains from this analysis and strains isolated from at the same sampling site one (Hunt *et al.*, 2008a), (Chapter 2, this thesis) to four years (Thompson *et al.*, 2005) prior were used in these assays: *V. breoganii* (1C10, ZF_29, ZF_55, FF_50), *Vibrio* sp. F10 (9ZB36,

9ZD137, 9CS118) and *V. crassostreae* (9ZC77, 9ZC40, 9ZC40, 9CS106). ZF_129 was initially used in competition assays, but was removed from subsequent analysis due to irregularities in its growth pattern. Particles used in *in vitro* assays were obtained by filtering 1000 L of water from same sampling site through a 64 μ M mesh net and sterilized by autoclaving them in seawater. Fifteen dead zooplankton and ten oblong plant-derived particles were picked by eye under a dissecting scope into 25 mls sterile seawater in 50 ml flasks. Each strain was grown in Luria Broth (BD Difco, Sparks, MD) plus 0.5 M NaCl for approximately 32 hrs. to an OD 600 nm of around 1.0. Twenty-five μ l of this culture from each competing strain was inoculated into 25 ml for the 1:1 samples. 10-fold or 100-fold dilutions were made in filter sterilized (0.2 μ M filtered, Pall, Ann Arbor, MI) spent media from the same strain and 25 μ l of these dilutions were used to make the 1:10, 10:1, 1:100 and 100:1 inoculation ratios with 25 μ l of undiluted culture of the competing strain. After two or four weeks at 22 °C with shaking, particles were picked with sterile forceps by eye under a dissecting scope, transferred into sterile seawater and put into 4 ml sterile seawater in a tissue grinder. These samples were ground and filtered into a 0.2 μ m sterivex filter (Millipore Corp., Billerica, MA) using a sterile lur-lock syringe. Tissue grinder tubes were washed with about 10 mls of sterile seawater and this was also transferred to the sterivex filter. Filters were placed at -20 °C immediately after filtering.

DNA Extraction

DNA was extracted using a modified method from Thompson et al 2004. Briefly,

sterivex filters were removed from the casing and cut using sterile razor in a sterile Petri dish. Entire, cut filters were placed in a 2 ml screw-cap tube with 0.2 g 0.1 mm zirconium beads (Biospec Products, Bartlesville, Okla.) with 750 µl cell lysis solution (PureGene Cell and Tissue Kit; Gentra Systems). Filters were bead- beaten at 5,000 rpm for 1 minute and placed at 80 °C for 5 minutes. These were often stored at 4 °C for up to one week. DNA extraction was done by incubating samples with 4 µl Rnase at 37 °C for 30 minutes, followed by addition of 250 µl protein precipitation solution (PureGene Cell and Tissue Kit; Gentra Systems). Samples were centrifuged for 3 minutes at 15,000 x g to remove filter debris, beads and precipitated protein and the supernatant transferred to a clean 1.5 ml eppendorf tube. Samples were spun again for 3 minutes at 15,000 x g and the supernatant was placed into a clean 1.5 ml eppendorf tube with 750 µl 100% IPA. Samples were centrifuged for 1 minute at 15,000 x g, and precipitated DNA was washed twice with 75% and 70% ethanol solution, dried for 30 minutes and resuspended in 100 µl DNA hydration solution (PureGene Cell and Tissue Kit; Gentra Systems) and stored at -20 °C. Genomic DNA extracted from overnight cultures for strains from each target population were done as previously described (Thompson *et al.*, 2005) to generate ROX labeled standards for peak identification. 9ZD137, 1C10, and 12B4 were used for as standards for peak identification of *Vibrio* sp. F10, *V. breoganii* and *V. crassostreae* populations, respectively.

Quantitative PCR

1-10 µl of extracted DNA from each sample was used as the template for quantitative

PCR as previously described (Thompson *et al.*, 2004). Briefly, samples were amplified for 25 to 30 cycles with unlabeled forward primer 567F, (5'-GGCGTAAAGCGCATGCAGGT-3') with 20 bp of the GC clamp (5'- CCCGCGCCCCCGTGCCCCCGCCCCGCGCCGGCCCCGGGCGCC-3') and reverse primers (680R, 5'-GAAATTCTACCCCCCTCTACAG-3') that target the *E. coli* region of the 16S between 576- 680 bp. PCR cycling conditions are as follows: 95 °C, 5 min; 30 cycles of 95 °C for 1 minute, 65 °C 2 minutes; 65 °C 5 minutes with Sure Start Taq (Stratagene). This was followed by 3 additional cycles of PCR using a fresh reagents and 5' Fluorescein amidite (FAM) labeled 567F primer with the full GC clamp (5'- GCCGCCTGCAGCCCGCGCCCCCGTGCCCCCGCCCCGCGCCGGCCCCGGGCGCC-3'), unlabeled reverse primer 680R and 2 µl of the previous PCR reaction as template. Standards used in peak identification were generated with the same cycling and labeling parameters as the FAM labeled samples, but using fluorescein 6 carboxy-X-rhodamine (ROX) labeled primers. For *Vibrio sp. F10*:*V. crassostreae* competition assays, although their target sequences have two bp differences between them, these peaks could not be resolved by differential melting under any cycling parameters used. Instead, 10 µl of labeled PCR product was added to 1 µl BsrI restriction enzyme (New England Biolabs, Ipswich, MA) and incubated at 65 °C for 1 hour. This restriction enzyme recognizes a cut site in the *V. crassostreae* amplified product that is missing from the *Vibrio sp. F10* product. The resulting fragments are easily separated, as the shorter *V. crassostreae* product travels faster through the capillary due to its length. This reaction mixture was loaded directly onto the capillary sequencer; other PCR products were diluted 1:50 in distilled water

before running on the capillary sequencer.

Constant Denaturing Capillary Electrophoresis and Peak Area Analysis

All labeled cut or uncut samples were injected onto a modified 96-well MegaBACE capillary sequencer (GE Healthcare) for 20-60 seconds along with ROX labeled peak standards (30 sec injection time) filled with linear polyacrylamine matrix (GE Healthcare). Samples were cycled between 54 °C and 57 °C for 30 sec each 20 times before dropping the temperature to 40 °C as the products crossed the laser detection area. Samples were run under the MegaBACE Instrument Control Manager Window (GE Healthcare) with the following standard options: Chemistry name, GT DyeSet2 (ET-ROX;FAM;NED;HEX); Chemistry application, Genotyping; Beam Splitter1, 540DRLP; Beam Splitter2, 570DRLP; Run Voltage, 7; Run time, 70; Injection voltage, 10; Temperature, 57; Lazer Index, 2; PMT Voltage1, 750; PMT Voltage2, 75. Sequence Analyzer 3.0 View and Edit (GE Healthcare) was used to export the profiles into text for further analysis. Peak area for each species was measured with AcqKnowledge 3.9 software (Biopac Systems, Goleta, CA).

Biolog and Growth Assays

Growth on Biolog GN2 plates (Biolog, Inc., Hayward, CA) was tested as follows: cells were grown as a lawn on TSB plus 1% salt and swabbed into marine inoculating media containing, 25 mg/ml NaCl, 8 mg/ml MgCl₂ and 0.5 mg/ml KCl to an optical density (600 nm) of 0.1. A suspension of 150 µl was inoculated into each well of the 96-well

plate (95 different carbon substrates plus water) and incubated at 22 °C for four days.

Optical density readings were taken for each well at 580 nm on 96-well plate Varioskan Flash Spectral Scanning Multimode Reader (Thermo Scientific, Rockford, IL).

Growth on alginate (alginic acid, Sigma), chitin (from crab shells, Sigma), cellulose (SigmaCell, Sigma), xylan (from birchwood, Sigma), carboxy-methyl-cellulose (Sigma) and hydroxy-ethyl-cellulose (Fluka Analytical) were done using a minimal media (Tibbles & Rawlings, 1994) without another source of carbon added to the media. All incubations were performed in 25 mm glass tubes for up to 4 weeks. OD readings were taken every week at 600 nm. Additionally, strains were tested for growth on alginate as sole carbon source using alginate slurry plates containing 15.7g alginic acid in 500 ml minimal media without a carbon source. Five µl of 2216 culture was spotted into four spots on the plate and growth was noted as dispersed colonies after 5 days.

Antagonistic interactions

Strain specific antagonistic interactions were identified as follows: cultures of all target and potential producer strains were grown in TSB plus 1% salt in 15 ml glass tubes for two days. 100 µl of the target strain was swabbed as a lawn onto brain heart infusion broth (Research Products International, Corp., Prospect, IL) plus 1% NaCl agar plates (BHI). Five µl of each potential producer strain was arrayed at regularly spaced intervals on the BHI plate. Interactions were noted after at least 24 hours and up to 1 week after streaking. Potential producer strains were marked as growing if they could be detected as a colony on the target lawn. Positive inhibition was noted when a zone of clearance was

observed around the potential producer strain colony. Generally, these inhibitions were still present after one week.

REFERENCES

- Acinas, S.G., Anton, J., & Rodriguez-Valera, F. (1999) Diversity of free-living and attached bacteria in offshore western Mediterranean waters as depicted by analysis of genes encoding 16S rRNA. *Applied and Environmental Microbiology*, **65**, 514-522.
- Bishop, J.K.B., Edmond, J.M., Ketten, D.R., Bacon, M.P., & Silker, W.B. (1977) Chemistry, Biology, and Vertical Flux of Particulate Matter from Upper 400 M of Equatorial Atlantic Ocean. *Deep-Sea Research*, **24**, 511-&.
- Crump, B.C., Armbrust, E.V., & Baross, J.A. (1999) Phylogenetic analysis of particle-attached and free-living bacterial communities in the Columbia river, its estuary, and the adjacent coastal ocean. *Applied and Environmental Microbiology*, **65**, 3192-3204.
- Delong, E.F., Franks, D.G., & Alldredge, A.L. (1993) Phylogenetic Diversity of Aggregate-Attached Vs Free-Living Marine Bacterial Assemblages. *Limnology and Oceanography*, **38**, 924-934.
- Ducklow, H.W. & Carlson, C.A. (1992) Oceanic Bacterial Production. *Advances in Microbial Ecology*, **12**, 113-181.
- Filippini, M., Buesing, N., Bettarel, Y., Sime-Ngando, T., & Gessner, M.O. (2006) Infection paradox: High abundance but low impact of freshwater benthic viruses. *Applied and Environmental Microbiology*, **72**, 4893-4898.
- Fraser, C., Alm, E.J., Polz, M.F., Spratt, B.G., & Hanage, W.P. (2009) The Bacterial Species Challenge: Making Sense of Genetic and Ecological Diversity. *Science*, **323**, 741-746.
- Gibbs, R.J. (1974) *Suspended solids in water* Pleum Press, New York.
- Goh, S.H., Potter, S., Wood, J.O., Hemmingsen, S.M., Reynolds, R.P., & Chow, A.W. (1996) WSP60 gene sequences as universal targets for microbial species identification: Studies with coagulase-negative staphylococci. *Journal of Clinical Microbiology*, **34**, 818-823.
- Guindon, S. & Gascuel, O. (2003) A simple, fast, and accurate algorithm to estimate large phylogenies by maximum likelihood. *Systematic Biology*, **52**, 696-704.
- Hubert, C., Loy, A., Nickel, M., Arnosti, C., Baranyi, C., Bruchert, V., Ferdelman, T., Finster, K., Christensen, F.M., de Rezende, J.R., Vandieken, V., & Jorgensen, B.B. (2009) A Constant Flux of Diverse Thermophilic Bacteria into the Cold Arctic Seabed. *Science*, **325**, 1541-1544.
- Hunt, D.E., David, L.A., Gevers, D., Preheim, S.P., Alm, E.J., & Polz, M.F. (2008a) Resource partitioning and sympatric differentiation among closely related bacterioplankton. *Science*, **320**, 1081-1085.

- Hunt, D.E., Gevers, D., Vahora, N.M., & Polz, M.F. (2008b) Conservation of the chitin utilization pathway in the Vibrionaceae. *Applied and Environmental Microbiology*, **74**, 44-51.
- Jeanmougin, F., Thompson, J.D., Gouy, M., Higgins, D.G., & Gibson, T.J. (1998) Multiple sequence alignment with Clustal x. *Trends in Biochemical Sciences*, **23**, 403-405.
- Lane, D.J. (1991). 16S/23S rRNA sequencing. In *Nucleic Acid Techniques in Bacterial Systematics* (eds E. Stackebrandt & M. Goodfellow), pp. 115-175. Wiley & Sons, Chinchester.
- Letunic, I. & Bork, P. (2007) Interactive Tree Of Life (iTOL): an online tool for phylogenetic tree display and annotation. *Bioinformatics*, **23**, 127-128.
- Long, R.A. & Azam, F. (2001) Antagonistic interactions among marine pelagic bacteria. *Applied and Environmental Microbiology*, **67**, 4975-4983.
- Martiny, J.B.H., Bohannan, B.J.M., Brown, J.H., Colwell, R.K., Fuhrman, J.A., Green, J.L., Horner-Devine, M.C., Kane, M., Krumins, J.A., Kuske, C.R., Morin, P.J., Naeem, S., Ovreas, L., Reysenbach, A.L., Smith, V.H., & Staley, J.T. (2006) Microbial biogeography: putting microorganisms on the map. *Nature Reviews Microbiology*, **4**, 102-112.
- Matz, C., McDougald, D., Moreno, A.M., Yung, P.Y., Yildiz, F.H., & Kjelleberg, S. (2005) Biofilm formation and phenotypic variation enhance predation-driven persistence of *Vibrio cholerae*. *Proceedings of the National Academy of Sciences of the United States of America*, **102**, 16819-16824.
- Nichols, W.W., Dorrington, S.M., Slack, M.P.E., & Walmsley, H.L. (1988) Inhibition of Tobramycin Diffusion by Binding to Alginate. *Antimicrobial Agents and Chemotherapy*, **32**, 518-523.
- Powell, J.H. & Meeuse, B.J.D. (1964) Laminarin in Some Phaeophyta of the Pacific Coast. *Economic Botany*, **18**, 164-166.
- Rosenzweig, R.F., Sharp, R.R., Treves, D.S., & Adams, J. (1994) Microbial Evolution in a Simple Unstructured Environment - Genetic Differentiation in *Escherichia-Coli*. *Genetics*, **137**, 903-917.
- Seymour, J.R., Ahmed, T., & Stocker, R. (2009) Bacterial chemotaxis towards the extracellular products of the toxic phytoplankton *Heterosigma akashiwo*. *Journal of Plankton Research*, **31**, 1557-1561.
- Simon, M., Grossart, H.P., Schweitzer, B., & Ploug, H. (2002) Microbial ecology of organic aggregates in aquatic ecosystems. *Aquatic Microbial Ecology*, **28**, 175-211.
- Sloan, W.T., Lunn, M., Woodcock, S., Head, I.M., Nee, S., & Curtis, T.P. (2006) Quantifying the roles of immigration and chance in shaping prokaryote community structure. *Environmental Microbiology*, **8**, 732-740.
- Tang, K.W., Freund, C.S., & Schweitzer, C.L. (2006) Occurrence of copepod carcasses in the lower Chesapeake Bay and their decomposition by ambient microbes. *Estuarine Coastal and Shelf Science*, **68**, 499-508.

- Thompson, J.R., Pacocha, S., Pharino, C., Klepac-Ceraj, V., Hunt, D.E., Benoit, J., Sarma-Rupavtarm, R., Distel, D.L., & Polz, M.F. (2005) Genotypic diversity within a natural coastal bacterioplankton population. *Science*, **307**, 1311-1313.
- Thompson, J.R., Randa, M.A., Marcelino, L.A., Tomita-Mitchell, A., Lim, E., & Polz, M.F. (2004) Diversity and dynamics of a north Atlantic coastal *Vibrio* community. *Applied and Environmental Microbiology*, **70**, 4103-4110.
- Tibbles, B.J. & Rawlings, D.E. (1994) Characterization of Nitrogen-Fixing Bacteria from a Temperate Salt-Marsh Lagoon, Including Isolates That Produce Ethane from Acetylene. *Microbial Ecology*, **27**, 65-80.
- Tilman, D. (1982) *Resource Competition and Community Structure* Princetown University Press, Princetown.
- Welch, R.A., Burland, V., Plunkett, G., Redford, P., Roesch, P., Rasko, D., Buckles, E.L., Liou, S.R., Boutin, A., Hackett, J., Stroud, D., Mayhew, G.F., Rose, D.J., Zhou, S., Schwartz, D.C., Perna, N.T., Mobley, H.L.T., Donnenberg, M.S., & Blattner, F.R. (2002) Extensive mosaic structure revealed by the complete genome sequence of uropathogenic *Escherichia coli*. *Proceedings of the National Academy of Sciences of the United States of America*, **99**, 17020-17024.

FIGURES AND TABLES

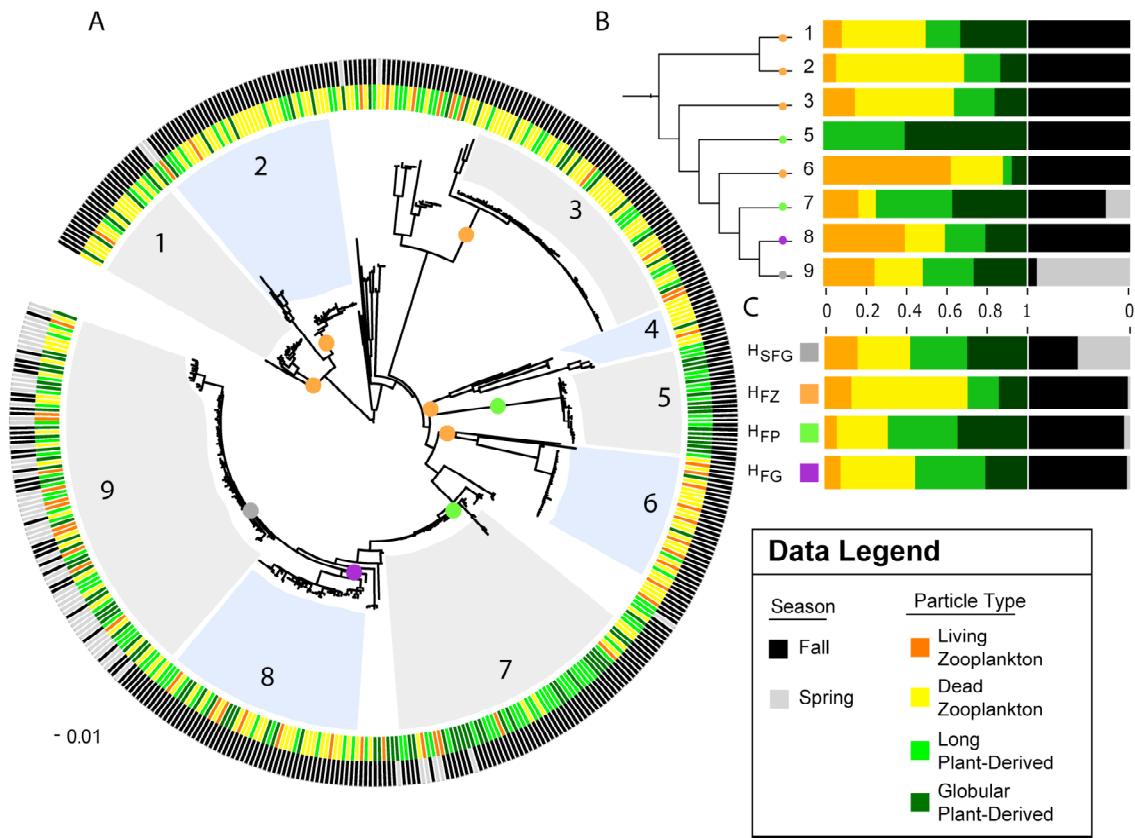


Fig. 1. Predicted populations, habitat preferences and environmental distributions for strains isolated from environmental sampling. **A.** Maximum likelihood phylogenetic tree derived from concatenation of partial *adk*, *mdh* and *hsp60* sequences for all strains (clones represented by one concatenated sequence) isolated from plant and zooplankton derived particles with rings designating season (outer ring) and particle type (inner ring) of origin for each strain. Circles at the nodes of each population are colored according to the predicted habitat preference as determined by the AdaptML algorithm(Hunt *et al.*, 2008a). Populations are alternatively shaded blue and gray and numbered corresponding to species as follows: 1, *Aliivibrio sp. I3*; 2, *Aliivibrio fischeri*; 3, *Enterovibrio calviensis*-

like; 4, *Enterovibrio norvegicus*; 5, *V. breoganii*; 6, *Vibrio sp. F10*; 7, *V. crassostreae*; 8, *V. tasmaniensis/V. lentus*; 9, *V. splendidus*. **B.** Normalized distribution (clones included) across sampled particle types for each of the predicted populations with more than 10 total representatives. **C.** Predicted habitats from AdaptML summarizing characteristic distributions across particles for populations. Scale bar is nucleotide substitutions per site.

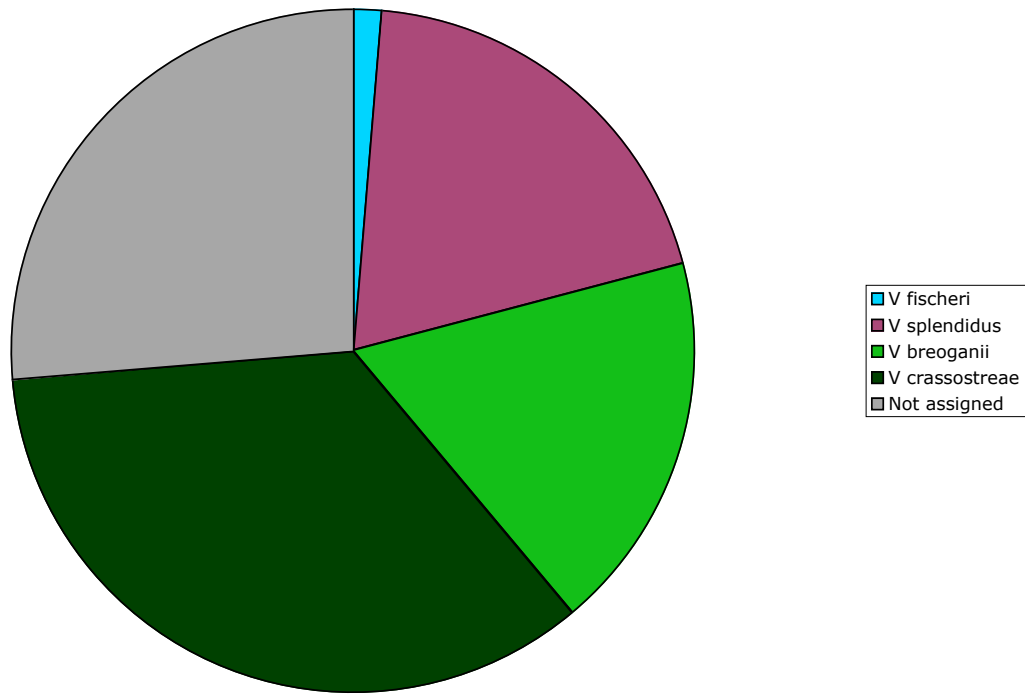


Fig. 2. *Vibrio* populations associated with brown algae (*Fucus vesiculosus*), one of the likely sources of plant-derived particles (n=72). The three most abundant populations on live brown algae also dominate plant-derived particles. Colors represent the proportion of each population out of all brown algae isolates across eight replicate samples. “Not

assigned” represent a diverse set *Vibrio* isolates that were not assigned to a significant population.

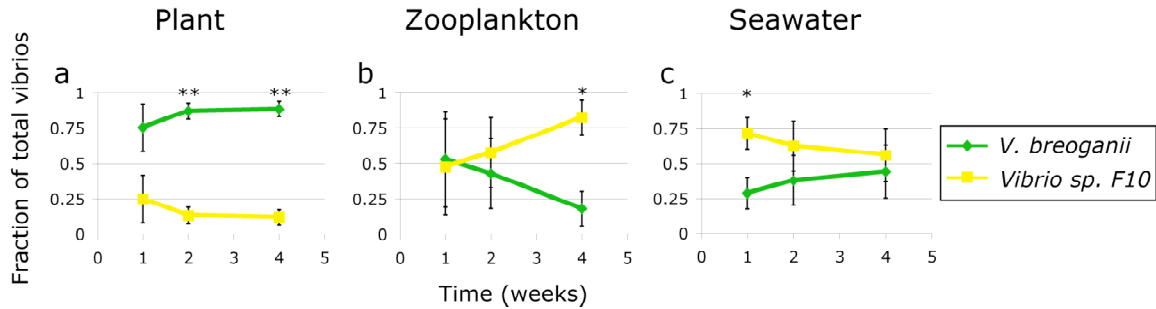


Fig. 3. Relative abundance of two strains with different predicted particle preferences on plant-derived (a) and zooplankton particles (b) or seawater alone (c). Plant and zooplankton particles were co-incubated in the same 25 ml flask, seawater inoculation was done separately. *Vibrio sp. F10* (yellow) has the strongest predicted preference for zooplankton, whereas *V. breoganii* (green) is one of two populations with a predicted preference for plant-derived particles. When testing one strain pair (1C10 and 9ZD137) in triplicate at equal inoculation ratios, the zooplankton-specialist population is reproducibly more abundant on zooplankton particles, and the plant- specialist is more relatively abundant on plant-derived particles, supporting their environmental distribution. (*) and (**) indicate statistical significance at the $p < 0.01$ and $p < 0.001$ respectively. Plant, plant-derived particles; Zooplankton, zooplankton particles; Seawater, incubation in sterile seawater alone (no particles added). Average of three replicate experiments, error bars represent standard deviation.

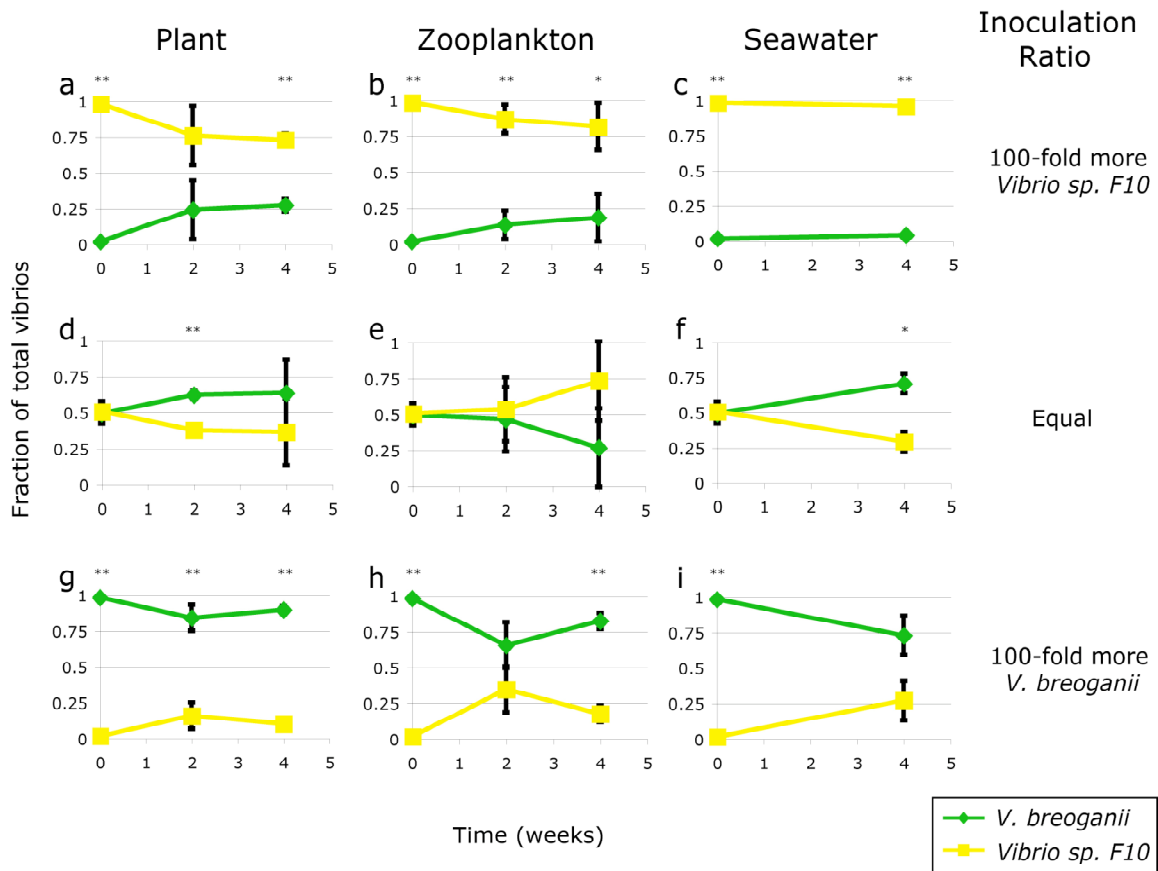


Fig. 4. Relative abundance of two populations with different predicted particle preferences on plant-derived (a, d, g) or zooplankton particles (b, e, h) or seawater alone (c, f, i). Plant and zooplankton particles were co-incubated in the same 25 ml flask, seawater inoculation was done separately. *Vibrio sp. F10* (yellow) has the strongest predicted preference for zooplankton, whereas *V. breoganii* (green) is one of two populations with a predicted preference for plant-derived particles. Here, replicate samples are derived from different strain pair combinations (biological replicates) and inoculation ratios varied from equal 100- fold in favor of each population. Despite higher average values of the specialist populations on their respective particles, particle

specialists are not statistically significantly more abundant on their respective particle type after four-week incubations. (*) and (**) indicate statistical significance at the $p < 0.01$ and $p < 0.001$ respectively. Plant, plant-derived particles; Zooplankton, zooplankton particles; Seawater, incubation in sterile seawater alone (no particles added). Average of three replicate experiments with different strains of the same species (biological replicates), error bars represent standard deviation. Time 0 values are measured from inoculation media, all other time points measured from particles or seawater.

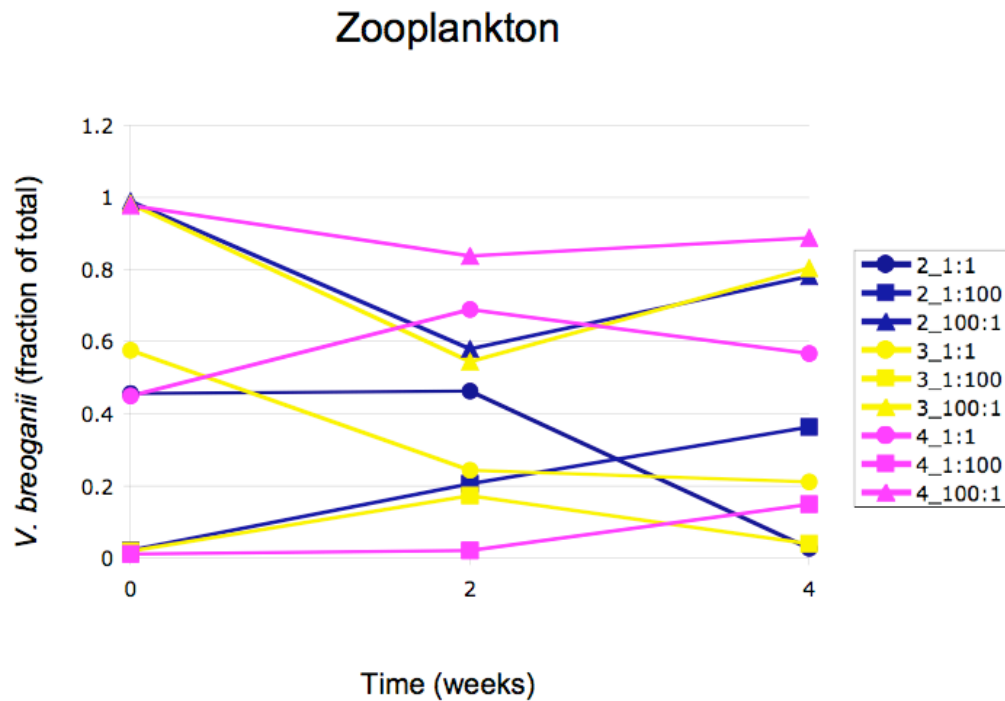


Fig. 5. Outcomes for individual strain-pair competition experiments with *Vibrio* sp. F10 (zooplankton specialist) and *V. breoganii* (plant-derived particle specialist) on zooplankton particles. After four- week incubations, initial starting density determines the relative abundance of each population for most replicates. Legend shows strain pair combination number followed by the initial inoculation ratio. Strain pair combinations 2, 3 and 4 are colored blue, yellow and pink, respectively correspond to the following strains (*V. breoganii* vs. *Vibrio* sp. F10): 2, strains 1C10 vs. 9ZD137; 3, ZF_29 vs 9ZB36; 4, FF_50 vs. 9CS118. Values for equal (1:1), or 100- fold more of either *V. breoganii* over *Vibrio* sp. F10 (100:1) or vice verse (1:100) are represented by circle, triangle, or square, respectively. Time 0 values are inoculation media ratios, all other time

points represent the fraction of *V. breoganii* as a fraction of total vibrios on zooplankton particles.

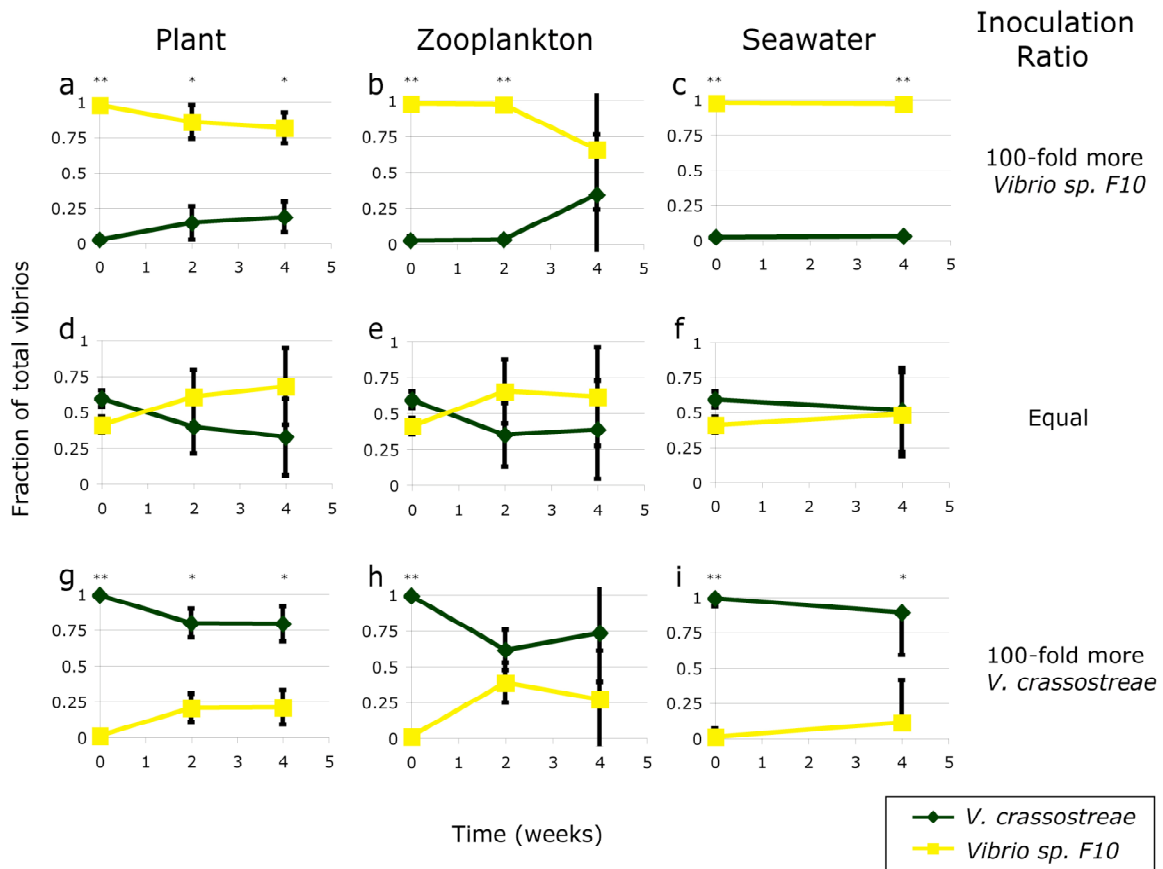


Fig. 6. Relative abundance of two populations with different predicted particle preferences on plant-derived (a,d,g) or zooplankton particles (b,e,h) or seawater alone (c,f,i). Plant and zooplankton particles were co-incubated in the same 25 ml flask, seawater inoculation was done separately. *Vibrio sp. F10* (yellow) has the strongest predicted preference for zooplankton, whereas *V. crassostreae* (dark green) is one of two populations with a predicted preference for plant-derived particles. Here, replicate samples are derived from different strain pair combinations (biological replicates) and inoculation ratios varied from equal 100- fold in favor of each population. Despite higher average values of the specialist populations on their respective particles, *Vibrio sp. F10* is

on average more abundant on all particle types, but this is not statistically significant. (*) and (**) indicate statistical significance at the $p < 0.01$ and $p < 0.001$ respectively. Plant, plant-derived particles; Zooplankton, zooplankton particles; Seawater, incubation in sterile seawater alone (no particles added). Average of three replicate experiments with different strains of the same species (biological replicates), error bars represent standard deviation. Time 0 values are measured from inoculation media, all other time points measured from particles or seawater.

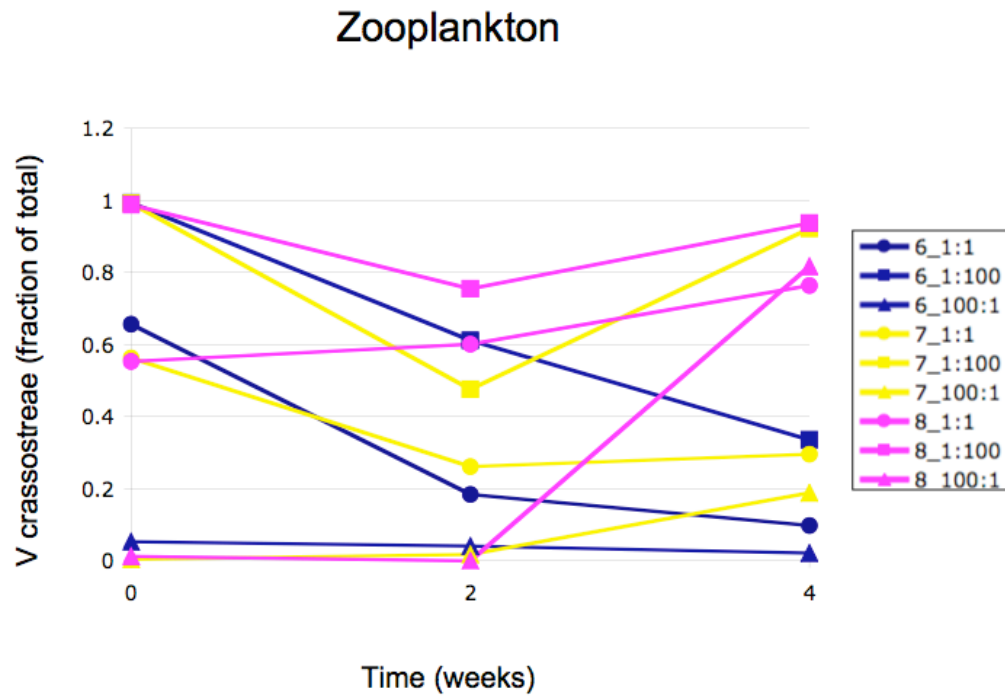


Fig. 7. Strain to strain variation in outcome in competition experiments with *Vibrio* sp. F10 (zooplankton specialist) and *V. crassostreae* (plant-derived particle specialist) on zooplankton particles. After four- week incubations, the difference in input ratios does not change the trend for two of the three biological replicates. Legend shows strain pair combination number followed by the initial inoculation ratio. Strain pair combinations 6, 7 and 8 are colored blue, yellow and pink, respectively correspond to the following strains (*Vibrio* sp. F10 vs. *V. crassostreae*): 6, strains 9ZD137 vs. 9ZC40; 7, 9ZB36 vs. 9ZC13; 8, 9CS118 vs 9CS106. Values for equal (1:1), or 100- fold more of either *V. crassostreae* over *Vibrio* sp. F10 (1:100) or vise verse (100:1) are represented by circle, square or triangle, respectively. Time 0 values are inoculation media ratios, all other time

points represent the fraction of *V. crassostreae* as a fraction of total vibrios on zooplankton particles.

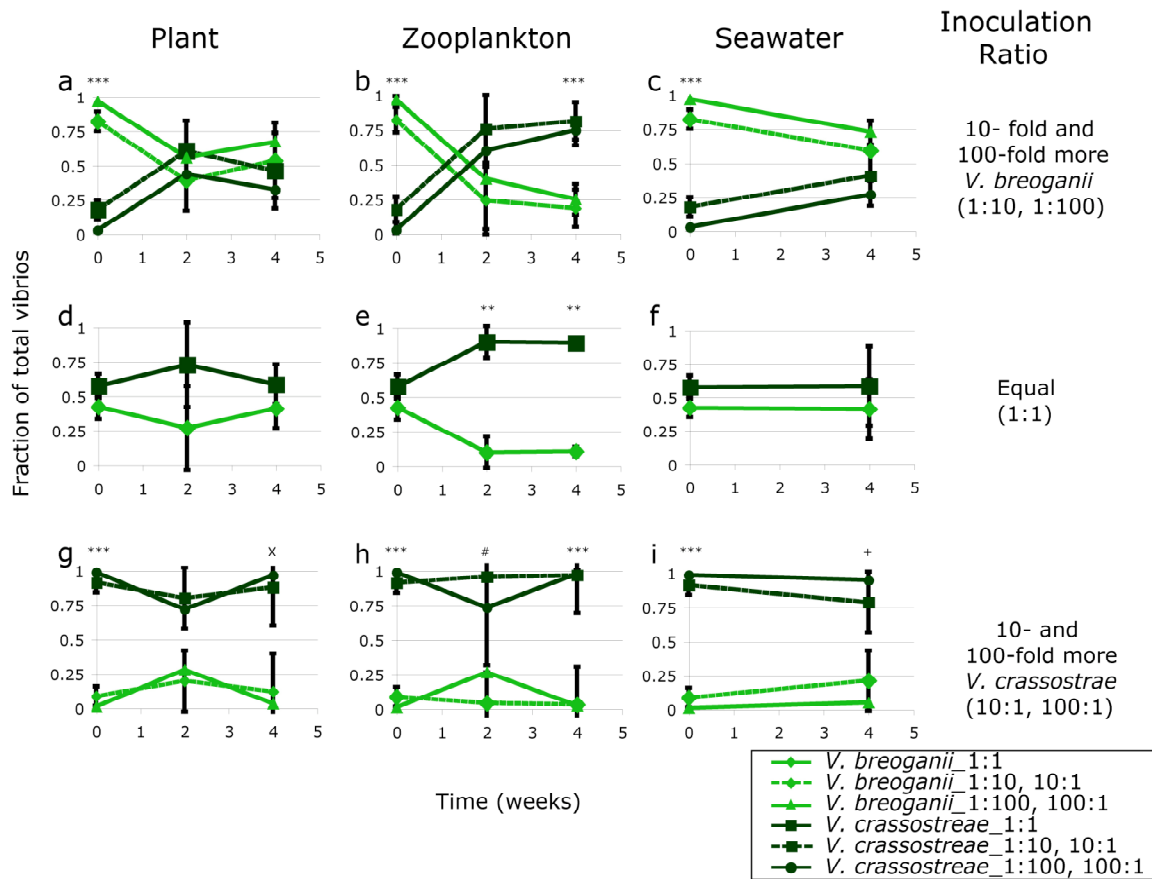


Fig. 8. Relative abundance of two populations with the same predicted particle preferences on plant-derived (a,d,g) or zooplankton particles (b,e,h) or seawater alone (c,f,i). Plant and zooplankton particles were co-incubated in the same 25 ml flask, seawater inoculation was done separately. Both populations have a predicted preference for plant-derived particles. *V. breoganii* (green) was not isolated from zooplankton, whereas *V. crassostreae* (dark green) was found on zooplankton at low relative abundances. Replicate samples are derived from different strain pair combinations (biological replicates) and inoculation ratios varied from equal 100- fold in favor of each population. *V. crassostreae* is statistically significantly more abundant than *V. breoganii*

on zooplankton particles, regardless of starting inoculation ratios. *V. crassostreae* is statistically more abundant on plant-derived particles after 4 weeks when inoculated at 10- and 100-fold higher concentrations than *V. breoganii*, but *V. breoganii* can not achieve statistically significantly higher values when inoculated at 100-fold higher abundance. Legends indicate species name followed by inoculation ratio. (**) indicate statistical significance at the $p < 0.001$ for the 1:1 incubations; (***) indicate statistical significance at the $p < 0.001$ for both 1:10 and 1:100 incubations; (x) indicate statistical significance of $p < 0.01$ and $p < 0.001$ for the 1:10 and 1:100 incubations, respectively; (#) indicate only the 1:10 incubation is statistically significant at $p < 0.001$. Plant, plant-derived particles; Zooplankton, zooplankton particles; Seawater, incubation in sterile seawater alone (no particles added). Average of three replicate experiments with different strains of the same species (biological replicates), error bars represent standard deviation. Time 0 values are measured from inoculation media, all other time points measured from particles or seawater.

Table 1. Differential growth on various carbon sources between particle type specialists.

	V. crassostreae				Vibrio sp. F10				V breoganii			
	9ZC77	9ZC13	9CS106	9ZC40	9ZB36	ZF_129	9CS118	9ZD137	FF_50	ZF_29	1C10	ZF_55
Tween 40	+*	++^	++	++	-	-	-	-	-	-	-	-
Tween 80	+	++	+	++	-	-	-	-	-	-	-	-
Dextrin	++	+++ [#]	++	++	+++	++	++	+++	+++	+++	+++	+++
D-Galactose	++	+++	+++	+++	-	-	-	-	+	+	+	++
Gentiobiose	-	-	-	-	-	-	-	-	+++	+++	+++	+++
D-Mannose	+	+++	+++	+++	-	-	-	-	-	-	-	-
D- Psicose	+	+	+	+	+	+	+	+	+	+	+	+
Sucrose	+	+++	+++	+++	+++	+++	+++	+++	++	-	-	-
D-Trehalose	+	+++	+++	+++	+++	+++	+++	+++	-	-	-	-
Acetic Acid	+	+	+	+	+	+	-	+	+	+	-	+
Cis-Aconitic Acid	+	+	-	+	-	-	-	-	-	-	-	-
Citric Acid	+	+	-	+	-	-	-	-	-	-	-	-
D-Gluconic Acid	+	+	+++	+++	-	-	-	-	+++	+++	+++	+++
alpha-Keto Glutaric Acid	+	++	++	++	-	-	-	-	-	-	-	-
D,L- Lactic Acid	+	+++	+++	+++	+	+	-	+	++	-	+	+
Succinic Acid	++	+++	++	++	-	-	-	-	+	-	-	-
Bromosuccinic Acid	+	+	++	++	-	-	-	-	-	+	+	-
D-Alanine	+	++	+	+	-	-	-	-	-	-	-	-
L-Alanine	+	+++	+++	+++	+	+	-	-	++	+++	+++	+++
L-Alanyl-glycine	++	+++	+++	+++	-	+	-	-	+++	+	-	+++
L-Asparagine	++	+++	+++	+++	+	+	-	-	-	+	+	++
L-Aspartic Acid	++	+++	+++	+++	+	-	+	-	-	-	-	+
Glycyl-L-Aspartic Acid	+	+++	+++	+++	-	+	-	-	+	+	-	++
Glycyl-L-Glutamic Acid	+	+++	++	+++	-	-	-	-	+	++	-	+
L-Histidine	+	+	+	++	-	-	-	-	-	-	-	-
L-Proline	+	++	++	++	-	-	-	-	-	-	-	+
L-Serine	++	+++	+++	+++	-	-	-	-	+	-	-	+
L-Threonine	+	+++	++	+++	-	-	-	-	+	++	++	+
Uridine	+	+++	+++	+++	+	+	+	-	+	+	+++	-
Thymidine	+	++	++	++	-	-	+	-	+	+	+++	-
Glycerol	+	+++	+++	+++	-	-	-	-	+	-	-	+
D-Glucose-6-Phosphate	+++	+++	+++	+++	-	-	-	-	-	-	-	-
Chitin	++	++	++	++	-	-	-	-	-	-	-	-
Alginate	+	+	+	+	-	-	-	-	+	+	+	+

*+ = growth as measured by OD600 between 0.15 and 0.5; cells colored light pink

^++=growth as measure by OD600 between 0.5 and 0.8; cells colored dark pink

#+++=growth as measured by OD600 greater than 0.8; cells colored red

Table 2. Distribution of plant-particle specialist populations across eight replicate samples of plant-derived particles.

Sample number	1	2	3	4	5	6	7	8
<i>V. breoganii</i>	14	1	1	14	0	0	1	1
<i>V. crassostreae</i>	10	8	2	10	12	9	7	4

Table 3. Antagonistic interactions between strains used in *in vitro* competition assays, including both growth characteristics and directed inhibition.

Target Strain		Potential Producer Strain											
		<i>Vibrio sp. F10</i>				<i>V. crassostreae</i>				<i>V. breoganii</i>			
		9CS118	9ZB36	ZF_129	9ZD137	9CS106	9ZC77	9ZC40	9ZC13	ZF_55	1C10	ZF_29	FF_50
<i>V. crassostreae</i>	9CS118	c ¹	c	c	c	c ³	c	c	c	c	c	c	c
	9ZB36	c	c	c	c	c	c	c	c	c	c	c	c
	ZF_129	c	c	c	c	c	c	c	c	c	c	c	c
	9ZD137	c	c	c	c	c	c	c	c	c	c	c	c
	9CS106	ci	c	c	c	c	c	c	ci ⁴	c	c	c	c
	9ZC77	c	c	c	c	c	c	c	c	c	c	c	c
	9ZC40	c	c	c	c	c	c	c	c	c	c	c	c
	9ZC13	c	c	c	c	c	c	c	c	c	c	c	c
	ZF_55	cm ²	cm	cm	cm	c	c	c	ci	c	cm	c	c
	1C10	ci	c	ci	c	c	c	c	ci	c	c	c	x ⁵
<i>V. breoganii</i>	ZF_29	cm	cm	cm	cm	c	c	c	ci	c	c	c	c
	FF_50	c	c	c	c	c	c	c	c	c	c	c	c

¹ c indicates colony formed, confirms growth of possible producer.

² m indicates small zone of clearing, possible producer slightly inhibits growth of target strain

³ larger, bolded, yellow letters represent interactions between strains used in competition assay. Of the three noted interactions between these strains, none resulted in dominance of potential producer regardless of initial density

⁴ i indicates large zone of clearing, possible producer strongly inhibits target strain

⁵ x indicates no evidence of growth of possible producer strain

Table 4. Temperature and Salinity Conditions During Spring and Fall Sampling Dates

Date	4/30	5/1	5/2	5/3	9/25	9/26	9/27	9/28
Water Temp (°C)	7	9	9.5	10	16	16	16.5	15
Salinity (ppt)	18-21*	25	21	25	32	32	32-33	33

*Range indicates salinity values at the beginning and the end of sampling.

Table 5. Distribution of habitat predictions over 100 replicate iterations of the clustering algorithm (AdaptML) representing relative habitat stability.

Total Number of Habitats	Percent of Trials
3	29
4	67
5	4

Table 6. Strain combinations for *in vitro* competition assays including initial inoculation

sources

Rep. No.	Ratio	Population 1	Strain 1	Source	Population 2	Strain 2	Source
1	1:1	<i>V. breoganii</i>	ZF_55	Culture	<i>Vibrio sp. F10</i>	ZF_129	Culture
1	1:100	<i>V. breoganii</i>	ZF_55	1:100 dilution	<i>Vibrio sp. F10</i>	ZF_129	Culture
1	100:1	<i>V. breoganii</i>	ZF_55	Culture	<i>Vibrio sp. F10</i>	ZF_129	1:100 dilution
2	1:1	<i>V. breoganii</i>	1C10	Culture	<i>Vibrio sp. F10</i>	9ZD137	Culture
2	1:100	<i>V. breoganii</i>	1C10	1:100 dilution	<i>Vibrio sp. F10</i>	9ZD137	Culture
2	100:1	<i>V. breoganii</i>	1C10	Culture	<i>Vibrio sp. F10</i>	9ZD137	1:100 dilution
3	1:1	<i>V. breoganii</i>	ZF_29	Culture	<i>Vibrio sp. F10</i>	9ZB36	Culture
3	1:100	<i>V. breoganii</i>	ZF_29	1:100 dilution	<i>Vibrio sp. F10</i>	9ZB36	Culture
3	100:1	<i>V. breoganii</i>	ZF_29	Culture	<i>Vibrio sp. F10</i>	9ZB36	1:100 dilution
4	1:1	<i>V. breoganii</i>	FF_50	Culture	<i>Vibrio sp. F10</i>	9CS118	Culture
4	1:100	<i>V. breoganii</i>	FF_50	1:100 dilution	<i>Vibrio sp. F10</i>	9CS118	Culture
4	100:1	<i>V. breoganii</i>	FF_50	Culture	<i>Vibrio sp. F10</i>	9CS118	1:100 dilution
5	1:1	<i>Vibrio sp. F10</i>	ZF_129	Culture	<i>V. crassostreae</i>	9ZC77	Culture
5	1:100	<i>Vibrio sp. F10</i>	ZF_129	1:100 dilution	<i>V. crassostreae</i>	9ZC77	Culture
5	100:1	<i>Vibrio sp. F10</i>	ZF_129	Culture	<i>V. crassostreae</i>	9ZC77	1:100 dilution
6	1:1	<i>Vibrio sp. F10</i>	9ZD137	Culture	<i>V. crassostreae</i>	9ZC40	Culture
6	1:100	<i>Vibrio sp. F10</i>	9ZD137	1:100 dilution	<i>V. crassostreae</i>	9ZC40	Culture
6	100:1	<i>Vibrio sp. F10</i>	9ZD137	Culture	<i>V. crassostreae</i>	9ZC40	1:100 dilution
7	1:1	<i>Vibrio sp. F10</i>	9ZB36	Culture	<i>V. crassostreae</i>	9ZC13	Culture
7	1:100	<i>Vibrio sp. F10</i>	9ZB36	1:100 dilution	<i>V. crassostreae</i>	9ZC13	Culture
7	100:1	<i>Vibrio sp. F10</i>	9ZB36	Culture	<i>V. crassostreae</i>	9ZC13	1:100 dilution
8	1:1	<i>Vibrio sp. F10</i>	9CS118	Culture	<i>V. crassostreae</i>	9CS106	Culture
8	1:100	<i>Vibrio sp. F10</i>	9CS118	1:100 dilution	<i>V. crassostreae</i>	9CS106	Culture
8	100:1	<i>Vibrio sp. F10</i>	9CS118	Culture	<i>V. crassostreae</i>	9CS106	1:100 dilution
9	1:1	<i>V. crassostreae</i>	9CS106	Culture	<i>V. breoganii</i>	ZF_55	Culture
9	1:10	<i>V. crassostreae</i>	9CS106	Culture	<i>V. breoganii</i>	ZF_55	1:10 dilution
9	1:100	<i>V. crassostreae</i>	9CS106	Culture	<i>V. breoganii</i>	ZF_55	1:100 dilution
9	10:1	<i>V. crassostreae</i>	9CS106	1:10 dilution	<i>V. breoganii</i>	ZF_55	Culture
9	100:1	<i>V. crassostreae</i>	9CS106	1:100 dilution	<i>V. breoganii</i>	ZF_55	Culture
10	1:1	<i>V. crassostreae</i>	9ZC77	Culture	<i>V. breoganii</i>	1C10	Culture
10	1:10	<i>V. crassostreae</i>	9ZC77	Culture	<i>V. breoganii</i>	1C10	1:10 dilution
10	1:100	<i>V. crassostreae</i>	9ZC77	Culture	<i>V. breoganii</i>	1C10	1:100 dilution
10	10:1	<i>V. crassostreae</i>	9ZC77	1:10 dilution	<i>V. breoganii</i>	1C10	Culture
10	100:1	<i>V. crassostreae</i>	9ZC77	1:100 dilution	<i>V. breoganii</i>	1C10	Culture
11	1:1	<i>V. crassostreae</i>	9ZC40	Culture	<i>V. breoganii</i>	ZF_29	Culture
11	1:10	<i>V. crassostreae</i>	9ZC40	Culture	<i>V. breoganii</i>	ZF_29	1:10 dilution
11	1:100	<i>V. crassostreae</i>	9ZC40	Culture	<i>V. breoganii</i>	ZF_29	1:100 dilution
11	10:1	<i>V. crassostreae</i>	9ZC40	1:10 dilution	<i>V. breoganii</i>	ZF_29	Culture
11	100:1	<i>V. crassostreae</i>	9ZC40	1:100 dilution	<i>V. breoganii</i>	ZF_29	Culture

12	1:1	<i>V. crassostreae</i>	9ZC13	Culture	<i>V. breoganii</i>	FF_50	Culture
12	1:10	<i>V. crassostreae</i>	9ZC13	Culture	<i>V. breoganii</i>	FF_50	1:10 dilution
12	1:100	<i>V. crassostreae</i>	9ZC13	Culture	<i>V. breoganii</i>	FF_50	1:100 dilution
12	10:1	<i>V. crassostreae</i>	9ZC13	1:10 dilution	<i>V. breoganii</i>	FF_50	Culture
12	100:1	<i>V. crassostreae</i>	9ZC13	1:100 dilution	<i>V. breoganii</i>	FF_50	Culture

Chapter Four

Reproducibility and Specificity of Vibrio Population Associations with Environmental Parameters³

³ To be submitted with co-authors Cutler, M. B., Timberlake, S., Alm, E. J. and Polz, M. F.

ABSTRACT

How reproducibly bacterial populations assemble under environmentally similar but spatially or temporally distinct conditions remains poorly understood. This is because few studies have attempted to define populations by repeatedly determining specific environmental associations by sampling broadly across a range of environmental parameters. Here we synthesize population structure data from three studies assessing the fine-scale distribution of *Vibrionaceae* among temporally and spatially distinct environmental categories in coastal seawater and animals. All studies use a dynamic model of ecological adaptation to identify bacterial population without relying on a pre-defined genetic cut-off or relation to previously named species. By comparing population predictions across studies, we show that a majority of populations display similar phylogenetic characteristics (identity and diversity), although on two occasions the model predicted multiple sub-populations. Additionally, most populations are congruent with previously characterized *Vibrio* species. Seven out of ten populations associated reproducibly with the same predicted habitat (large particles in the water column). Within the *Vibrionaceae*, specialization as evidenced by limited association with sampled environmental categories is rare but does occur. The meta-analysis provides a solid foundation with which to investigate the environmental factors that determine the relative abundances of populations in the environment.

INTRODUCTION

Vast taxonomic diversity has been observed among both environmental (Huber *et al.*, 2007; Thompson *et al.*, 2005b; Venter *et al.*, 2004) and animal-associated microbes (Eckburg *et al.*, 2005; Ley *et al.*, 2006). However, for the majority of these organisms, it remains unknown what ecological role they play or which environmental niche they occupy. The challenge is to identify groups of related organism, which act as ecologically cohesive populations (Hunt *et al.*, 2008a; Koeppel *et al.*, 2008). Additionally, identifying where organisms are abundant and why offers insight into mechanisms that generate and maintain diversity, including speciation and dispersal (Martiny *et al.*, 2006). This requires knowledge of the specificity and reproducibility of associations between such groups of organisms and environmental parameters that define their habitat.

Information on the ecological role of bacterial populations can be gained from identifying associations of genotypes with chemical (Lozupone & Knight, 2005) or physical gradients (Johnson *et al.*, 2006), or associations with other organisms (Visick & McFall-Ngai, 2000). However, in some cases, population structure may arise from efficient dispersal so that presence may not provide information on the ecological role, as may be the case for organisms that experience high dispersal into areas of the environment that do not support their growth (Hubert *et al.*, 2009). Alternatively, the effect of limited dispersion may locally create population bottlenecks, which result an association with an environmental categories as the result of stochastic processes (Fraser *et al.*, 2009). Determining the relationship of bacterial populations and environmental

parameters over multiple time points can help identify consistent patterns of association and potentially rule out spurious associations by local bottlenecks. Subsequently, putative populations exhibiting patterns of specialization can be investigated for genetic or genomic signatures reflecting this specialization, such as specific metabolic capabilities. Although many studies have correlated genotypic information (e.g., marker gene sequences) with environmental parameters to gain insight into the role of specific microbial populations in the environment [e.g., (Broughton & Gross, 2000; Furlong *et al.*, 2002; Lozupone & Knight, 2005)], many more questions can be answered once the reproducibility and specificity of the associations is known.

Here we analyze three studies investigating fine-scale population structure in the *Vibrionaceae* to determine the reproducibility and specificity of associations in the environment over time. All studies used a method of identifying bacterial populations that deviates substantially from previous analyses of genetic diversity. Typically, the vast microbial diversity contained within environmental samples is categorized using an arbitrarily defined cut-off value of genetic similarity, typically using 16S rRNA nucleotide identity (Huber *et al.*, 2007; Li *et al.*, 2009). In contrast, Hunt, David *et. al.* (2008) created a dynamic model that groups bacteria by both genetic and ecological similarity (AdaptML). Using this approach, twenty-five populations were identified as displaying varying degrees of being free-living or associated with different sized-particles in coastal water across two seasons (this study will subsequently be referred to as fract-2006). Additional sampling at the same site occurred in both the spring and fall of the following year and the same method of analysis was used to identify bacterial

populations. Chapter 2 of this thesis extends the analysis to include associations with the digestive and respiratory tract of marine invertebrates (subsequently referred to as invert-2007). Finally, Chapter 3 of this thesis further investigates associations with different particles types contained within the largest sized particle fraction from fract-2006 (subsequently referred to as particle-2007). Although there were some methodological differences between studies, which are summarized in Table 1, comparing the results from these three studies provides valuable information of how reproducible and specific environmental associations are within the *Vibrionaceae*.

While some attempts have been made to organize prokaryotic diversity using evolutionary (Koeppel *et al.*, 2008) or ecological (Hunt *et al.*, 2008a) information, these methods have yet to be incorporated into standard taxonomy. Here we also attempt to answer how compatible taxonomy and a more natural organization of bacterial diversity might be. The *Vibrionaceae* represent a good model as they have been subject to extensive taxonomic inquiry and recent revision using multilocus sequence analysis (Le Roux *et al.*, 2004; Sawabe *et al.*, 2007; Thompson *et al.*, 2005a), which has the potential to identify environmentally relevant genotypic clusters (Gevers *et al.*, 2005). By comparing populations modeled according to ecological association to previously named species and to populations predicted across similar studies, resulting similarities (and differences) between modeled populations and previously named species may encourage adaptation of a species concept for microbiology that is grounded in ecological and evolutionary theory.

Our model, the *Vibrionaceae*, is a cohesive family of gram negative gamma-proteobacteria sharing the rare feature of two circular chromosomes, which have been evolving together throughout the history of the *Vibrionaceae* (Okada *et al.*, 2005). Many members of this family are pathogenic to both humans and marine biota and *Vibrio* species are easily isolated from a range of coastal environments (Thompson *et al.*, 2004a). The conserved chitin utilization pathway across *Vibrio* species makes this family important in nutrient cycling, as degraders of the highly abundant marine polymer, chitin (Hunt *et al.*, 2008b). All of these features lend support for this bacterial family as a good model system with which to explore co-existence and reproducibility of bacterial populations.

RESULTS AND DISCUSSION

Population Identification: Equivalent Populations

Populations in each study were defined by a model of ecological differentiation without a pre-determined genetic similarity cut-off and without knowledge of existing species or other populations. Thus, predicted populations may extend across current species boundaries, or divide what is currently considered one species. Additionally, the genetic diversity of populations predicted across studies need not be similar. The relationship of populations predicted across studies and of populations with named *Vibrio* species is assessed using phylogeny of one to five housekeeping genes.

A unique feature of the algorithm is that it determines how much of the total diversity should be contained within each population (genetic breadth). First, the algorithm conservatively estimates the number of true habitats (or niches) from the combined phylogenetic and environmental data. Next, strains are organized into populations based on their predicted habitat and genetic similarity. The genetic breadth of a population is defined as broadly as possible, such that all strains have the same habitat prediction. Finally, low confidence populations are filtered out through empirical statistical testing. These low confidence populations were left out of subsequent analyses with the exception of two cases: In both invert-2007 and particle-2007, a large group of strains with the same habitat prediction failed to pass the significance test because the distribution of these strains was nearly randomly distributed across all environmental parameters. These “generalists” were retained as a population because they comprised a significant proportion of the total isolates in each study. In proceeding sections, the “generalists” will be included in the population analysis. Strains in other low confidence populations will be included in this analysis if they fall within the subtree of a population predicted in another study.

To determine how populations identified across all studies are phylogenetically related, a maximum likelihood tree was constructed from partial sequences of a housekeeping gene (*hsp60*) from all strains used in each study (Fig. 1). Populations are identified as equivalent if the subtree containing all isolates from the population in one study overlapped with most or all isolates from a population in another study. In many cases, the genetic breadth of equivalent populations across studies is in good agreement.

V. breoganii (F9) is a good example of this since this population is essentially identical in genetic breadth in all studies (Fig. 1). However, occasionally the genetic breadth of a population in one study extends beyond that of an equivalent population in another study. When this occurs, either isolates are included that were not part of the equivalent population, or the genetic breadth of a population in one study overlaps with multiple populations in the other study (Fig. 1, Population *V. tasmaniensis*- P8). If the extended genetic breadth results from the inclusion of only a few additional strains, the populations are considered equivalent despite this discrepancy, since the majority of isolates overlap. Otherwise, the population with the largest genetic breadth determines the extent of the population, and other populations contained therein are considered sub-populations. In this way, equivalent populations and sub-populations can be used in further analysis.

While the diversity of predicted populations across studies is relatively consistent, there is no way to assess the stability of sub-populations since the experimental design, which identified subpopulations in one study, was not repeated in any of the other studies. Within the *V. tasmaniensis* population (P8), multiple sub-populations were predicted in invert-2007 (Fig. 1). Because the environmental categories sampled in invert-2007 were largely unique to that study, it is impossible to determine whether the model would predict similar diversity for sub-populations if the experiment were repeated. A similar situation occurred within the *V. splendidus* (I16) population, where fract-2006 predicted multiple sub-populations within this group. It will be interesting to determine whether and how many of these sub-populations represent stable lineages evolving independently from the overall population.

Overall, thirteen populations were identified; only two of these populations contained sub-populations, where multiple populations were identified within the genetic bounds of a population predicted from an alternative study. This suggests that in general, populations display similar amounts of ecological differentiation across a range of environmental categories.

Population Identification: Named Species

Matching populations to previously named species provides an opportunity for additional insight from species characteristics. Once populations are given species names, information from previous research, including the properties determined during initial characterization of the species, help to identify how these organisms may be adapted to conditions in their preferred habitat. However, populations may also prove to be genetically distinct from previously named species, warranting classification as a novel species. As the current criteria to define species in microbiology lacks ecological or evolutionary justification, merging populations predicted from ecological principles with what is currently defined as a species will lay the foundation for an improved assessment of the species concept in microbiology.

Closely related *Vibrio* species can often be discriminated genetically by sequencing multiple housekeeping genes. Additionally, the use of multiple genes provides a more robust picture of phylogeny given high frequency of homologous recombination in bacteria (Vos & Didelot, 2009). To this end, populations were compared with named bacterial species at loci widely used in phylogenetic classification within the vibrios: DNA gyrase B (*gyrB*) (Le Roux *et al.*, 2004; Sawabe *et al.*, 2007),

RecA/RadA recombinase (*recA*) (Thompson *et al.*, 2005a), DNA topoisomerase I (*topA*) (Sawabe *et al.*, 2007), uridylate kinase (*pyrH*) (Thompson *et al.*, 2005a) and ATP synthase F1, alpha subunit (*atpA*) (Thompson *et al.*, 2007). These five genes provide the genetic resolution to identify most populations as members of named species.

Four populations are clearly distinct from previously named *Vibrio* species based on phylogenetic position at more than one housekeeping gene. Of these populations, two show a clear relationship to named *Vibrio* species. These are always the closest relatives of a single named species, but fail to cluster within the genetic breadth of that named species. These populations have been provisionally named *Enterovibrio calviensis*-like (F1) and *V. rumoiensis*-like (F4), to represent their relationship to these two names species (Fig. 2-3, Fig. 5-6). Two other populations groups did not correspond to previously named species and seem to be distinct from any known vibrios. Their phylogenetic relationship within the *Vibrionaceae* changes with each genetic locus used. Population F10, with consistent population predictions in all studies, is not closely related to any named *Vibrio* species based on three housekeeping genes and 16S rRNA nucleotide identity. Its closest relatives based on 16S rRNA marker gene similarity is *V. gallaecus* (96% best Blast hit). However, in other loci, this population is closest to *V. gazogenes* at *atpA* (Fig. 2) and *V. kanoleae* at *gyrB* (Fig. 3). The relationship with other species at all genetic loci is summarized in Table 2. Population F6, making up a small percentage of the total isolates in only two studies, is also not closely related to any named *Vibrio* species. It displays a similar pattern of mixed phylogenetic signals (Table 2). Populations F6 and F10 have been given the provisional names *Vibrio sp.* F6 and

Vibrio sp. F10. Further tests are needed to determine whether distinct populations should be further characterized as a novel species within the *Vibrionaceae*.

Most other populations correspond to previously named species. Several species have been characterized, which are genetically very similar to *V. splendidus* (Le Roux *et al.*, 2004; Macian *et al.*, 2001; Thompson *et al.*, 2003a; Thompson *et al.*, 2003b) but these are typically distinguishable by only a few of the commonly sequenced housekeeping genes and recombination at some loci may confuse the phylogenetic signal (Sawabe *et al.*, 2007). Species assignments could be made for all populations closely related to *V. splendidus*, although some conflicting signals were observed. Sub-populations F18, F19, F23 and F25 (from fract-2006) grouped with *V. splendidus* type strains (Figure 2-6). However, any ecological differentiation occurring within these sub-populations has not yet created genetic differentiation at most genetic loci used, highlighting the difficulty of using genotypic information alone to identify ecologically distinct populations. Population F11 is differentiated from *V. splendidus* at *hsp60* (Fig. 1), *pyrH* (Fig. 4) and *topA* (Fig. 6) loci, but otherwise grouped with the same *V. splendidus* type strains above. Thus, sub-population F11 displays some signs of (possibly beginning) genetic differentiation. Whether these genetic differences and corresponding ecological differentiation are enough to warrant distinct species classification remains to be seen. Other populations in addition to those closely related to *V. splendidus* correspond to named species at multiple different housekeeping loci (Table 2), providing strong phylogenetic support that both the genetic identity and breadth of populations predicted

by AdaptML correspond in large measure with named *Vibrio* species. Populations will subsequently be referred to by species names whenever possible according to Table 2.

Reproducible Associations with Environmental Parameters

In both fract-2006 and particle-2007, vibrios were isolated from particles greater than 64 μm in size. Thus comparing populations associated with these particles can provide a measure of how reproducible the associations are with particles of this type although methodological differences between these studies may create some differences in predicted associations. Most importantly, in fract-2006, isolates were obtained from all >64 μm particles collected as a single fraction by filtration, while in particle-2007, strains were isolated non-exhaustively from visually identifiable particles, hand picked under a dissecting microscope. Moreover, studies were conducted at slightly different times of the year when seawater temperatures (Table 1) and other environmental factors were differed to some degree. Also, fract-2006 aimed at isolating more strains from each of the four replicate samples obtained on the same day, while invert-2007 and particle-2007 isolated fewer strains from each sample but sampled more broadly across eight replicates on multiple days. Finally, populations comprising only a small percentage of total isolates across all studies might show variable association with these particles due to random sampling. To this end, populations that did not comprise more than 2% of the total within any study are not considered all proceeding analyses.

Despite methodological differences, associations with >64 μm particles were reproducible for seven of the ten populations that were predicted from the two studies.

Discrepancies are defined as more than 6% difference between the percent of isolates obtained from >64 µm particles (as a percentage of total isolates from >64 µm particles) within equivalent populations across studies (Table 3). Additionally, AdaptML predicted different habitats for two of three discordant populations. The model predicted an alternate primary habitat for *V. tasmaniensis* (small particles between 5 µm-64 µm) and *Enterovibrio calviensis*-like (free-living) in fract-2006, whereas they were both predicted as a significant population associated with >64 µm particles in particle-2007. Variability in host abundance may have caused the discrepancy between these studies.

The model suggested reproducible association with >64 µm particles in both studies for six out of the seven remaining populations. *V. cyclitrophicus* is the only population that was predicted to be associated with the >64 µm fraction from fract-2006 but was not predicted to be a significant population in particle-2007. Because the relative abundance of this population associated with the >64 µm particle is greater in fract-2006 than particle-2007, it may be associated with a particle type not sampled in the 2007 analysis.

Population Specialization

Populations with highly specialized environmental associations can be determined using data across all three studies. Populations with very specialized environmental associations are likely to possess relative growth or survival advantage in a specific environmental habitat. Widely distributed populations could have efficient dispersal mechanisms rather than specific growth capabilities, or have a capacity for growth on a

wide range of substances. Differences between specialized populations and those more widely distributed may inform future studies on the genetics of specialization.

Populations were grouped by correlated environmental distributions using an agglomerative hierarchical processing method (Eisen *et al.*, 1998). Populations with highly similar distributions over all categories sampled in the three studies are grouped into the same profile. Populations within each of the resulting profiles types (Fig. 7, I-IV), share many characteristics. Taxonomic groups with type IV profiles are relatively more abundant in the spring. Populations within the type II profile frequently are associated with animals. *Aliivibrio fischeri*, which is known to be the symbiont of the light organ of squids (Visick & McFall-Ngai, 2000) has the type II profile, along with populations that tend to have higher abundances in crabs, mussels and on >64 μm particles. Populations with type I environmental profiles are associated with >64 μm particles, but have lower abundances in crab and mussels. Finally, populations with type III profile tend to have lower abundances in crab and mussel but higher abundances on small particles and as free-living cells. Two out of these four profiles (type I and III), have the most restricted overall environmental distributions.

Two populations were both abundantly and reproducibly associated with a specific environmental parameter, making a strong case for environmental specialization. *V. ordalii* is reproducibly present in the free-living state in the water column, since it was predicted as free-living in fract-2006 and it was also found at high relative abundances in seawater in invert-2007 (Fig. 7). Although it was also found in mussels (invert-2007) and in the 1 μm fraction (fract-2006), these abundances were very low compared to the

relative abundance in the free-living state and may result from accumulation in mussels due to filter-feeding or large cells ($>1\ \mu\text{m}$) being present in the water, respectively. Another group that displayed a reproducible association with a specific habitat was *Vibrio* sp. F10. It was found to be associated with $>64\ \mu\text{m}$ particles in fract-2006, and specifically with zooplankton in both invert-2007 and particle-2007. It was also found at low relative abundances in the crab and mussel, but these values were low compared to the abundances in the zooplankton. This specificity was demonstrated by habitat predictions across all three studies, making this association specific and robust. Although *Enterovibrio norvegicus* and *Aliivibrio logei* seem to have a specialized association with small particles and crabs, respectively (Fig. 7), these environmental parameters were not re-sampled, so it remains to be seen whether their specificity results from a reproducible association. This demonstrates a high degree of reproducibility of population association with sampled environmental categories for these populations.

Other populations displayed more generalist distribution across all sampled categories. Whether distributions for these more generalist populations are dominated by the effects of dispersal or high relative growth and/or survival rates in all sampled categories has yet to be determined. *V. splendidus* displays a highly generalist distribution across sampled categories (Fig. 7), but this is possibly due to their ability to grow at low temperatures more efficiently than other populations. It is possible that the same mechanism of temperature adaptation promotes their relative growth over other populations in almost all environmental areas, making them widely distributed across all environmental categories. Early colonization may result in persistence once other

populations gain a growth advantage as temperatures increase, depending on the ability of populations to invade into the pre-colonized area.

Merging ecological data from multiple studies can also provide information about populations that have a similar or correlated pattern of occurrence. Genetically similar taxonomic groups are expected in similar environments since they share many of the same properties (environmental filtering); however, their overlapping niches may cause similar populations to be excluded from the same environment through competitive exclusion (phylogenetic overdispersion) (Webb *et al.*, 2002). As there are four closely related populations within the *V-splendidus* clade (*V. crassostreae*, F13; *V. tasmaniensis*, P8; *V. cyclitrophicus*, F15; *V. splendidus*, I16) and three closely related populations in the *Aliivibrio fischeri* – *Aliivibrio logei* clade (*Aliivibrio fischeri*, F8; *Aliivibrio logei*, F7; *Aliivibrio sp.* I3), these are appropriate populations with which to address this question (Fig. 1A).

The environmental distribution of many of these genetically similar populations are more similar to populations that are not members of the same clade (Fig. 7), demonstrating that genetic similarity does not predict environmental similarity in these populations. *Aliivibrio sp.* I3, *Aliivibrio fischeri* and *Aliivibrio logei* have environmental profile types I, II and IV, respectively, and are each correlated with non-*Aliivibrio fischeri* like populations. Within the *V. splendidus*-clade, *V. crassostreae*, *V. tasmaniensis* and *V. cyclitrophicus* all have the same environmental profile type (III), but *V. crassostreae* is highly correlated with *V. breoganii*, which is not a member of this clade. The two clusters (F10 and I16) of *V. splendidus* are more similar to each other than to

other populations, but they are distinct from the other members of the *V. splendidus*-clade.

This overall structure suggests that competition between populations, such as for nutrients on chitinous or algal particles, dominate the interactions between these groups, rather than phenotypic attraction, which would tend to cluster these groups in the same environment. Horner-Devine and Bohannan (Horner-Devine & Bohannan, 2006) quantitatively analyzed the genetic diversity of bacterial communities by comparing multiple studies for evidence of environmental filtering or competitive exclusion. Their analysis suggest that competition does not play an overwhelming role in shaping bacterial communities because communities are made up of taxa that are commonly more closely related than expected by chance. However, they admit that many of their analyses include so many different populations (spanning all Alpha- and Beta- proteobacteria and Cytophaga-Bacteroides-Flavobacteria) that it may be difficult to quantitatively detect the effects of competitive exclusion between close relatives which results in phylogenetic overdispersion. Spatial scale is also important, since organisms need to physically co-exist to interact. Thus, confining the sampling to a small area will reduce the chance of sampling across multiple true habitats. While the effects of competitive exclusion may be difficult to detect over broad phylogenetic distances and spatial scales, both the genetic breadth and the scale of sampling used across these studies seem to be sufficient to detect such signals.

CONCLUSIONS

In this meta-analysis, bacterial strains collected from different environmental habitats at the same sampling site over two years are used to refine information on *Vibrio* population structure in the coastal environment. Reproducible association with large particles in the water column was seen for seven out of ten populations, suggesting that *Vibrios* have a robust and predictable connection to these particles. A high degree of specialization, as evidenced by population associations with a limited number of environmental categories, is rare and could only be reproducibly demonstrated for two out of thirteen populations. Efficient dispersal mechanisms may account for most populations being isolated from a wide range of environmental categories, along with retaining a large metabolic capacity for growth on a variety of substrates. Additionally, populations predicted by a model of ecological adaptation across studies had similar genetic breadth and often corresponded to previously named species within the *Vibrionaceae*. Two populations were divided into multiple sub-population by the model, but it remains to be seen whether these populations are reproducibly predicted by the model when the experiment is repeated.

Determining reproducibility and specificity of microbial taxonomic groups will provide a better understanding of how bacteria interact and evolve. More efforts should be made to identify ecologically relevant bacterial populations, adding to 16S rRNA gene-based identification with more discriminating markers that identify more closely related (an more recently differentiated) populations beginning to evolve independently. Cultivation is a feasible mechanism of study for this family of bacteria, but other

methods, such as single-cell amplification of multi-locus genes or single-cell genomics, may be needed to study environmentally important bacterial families that are difficult to culture. Sampling environmental categories at the microscale, the scale appropriate for bacteria, can be a challenge when studying bacterial population structure, and must be validated using both experimentation and reproducibility. However, if bacterial populations can be shown to have a reproducible association with environmental categories, these associations can be used to understand ecological factors that drive their evolution.

MATERIALS AND METHODS

DNA Extraction and Illumina Sequencing

Representative strains of most populations were chosen for additional multi-locus and Illumina sequencing, beyond what had been done in the original studies. Cultures of each of these representative strains were grown in tryptic soy broth plus 2% NaCl. DNA was extracted from 100 ml of two- day culture using Qiagen Genomic Buffer set (cat# 19060) and the Genomic Tip 500 (cat# 10262) following the manufacture's protocol (Qiagen, Valencia, CA). Samples were sent to the Broad Institute (MIT, Cambridge, MA) for sequencing-by-synthesis principle (Kaller *et al.*, 2007) based on Illumina technology (San Deigo, CA).

Additional Multi-locus Sequences for Population Representative

Primers 27f and 1492r (Lane, 1991b) were used to amplify the small subunit (16S) rDNA sequence using DNA extracted for Illumina sequencing. Additionally, partial *gyrB* and *recA* sequences were generated using the following primer sets: *gyrB*_Vfmod.for (5'-CGT TTY TGG CCR AGT G -3') *gyrB*.rev (5'- TCM CCY TCC ACW ATG TA -3'), *recA*.for (5'- TGG ACG AGA ATA AAC AGA AGG C -3') and *recA*.rev (5'- CCG TTA TAG CTG TAC CAA GCG CCC -3'). *gyrB* and *recA* genes were amplified using the following PCR conditions: 95 °C for 3 min.; 30 cycles of 95 °C for 30 sec, 40-45 °C for 30 sec and 72 °C for 1 min.; 72 °C for 5 minutes (annealing temp for *gyrB*: 40 °C, *recA*: 45 °C). All genes were sequenced in one direction only. Sequencing was performed at the Bay Paul Center at the Marine Biological Laboratories in Woods Hole. Automatic base calls were trimmed and manually curated using Sequencher (Gene Codes Corp., Ann Arbor, MI) and aligned using Clustalw (Thompson *et al.*, 1997) with visualization and further manual curation using MacClade (Sinauer Associates, Sunderland, MA).

Some *gyrB* and *recA* and all of the *atpA*, *topA* and *pyrH* sequences for population representatives were obtained by assembling fragments of the genomes obtain by Illumina sequencing against a reference multi-locus gene sequence. We used maq (version 0.7.1) to map the single-end Illumina reads to the nucleotide coding sequence of a related strain (Table 4). Maq's default parameter set was modified in the following ways to allow for the larger-than-usual divergence expected between reads and reference:

(i) Mapping was altered to search for three mismatches in the seed (the maximum

allowable) and 20 mismatches in the full sequence. The threshold on the sum of mismatching base qualities was increased to 300 (to allow for high quality mismatches expected due to divergence, as well as many low-quality mismatches due to the relatively long reads). (ii) In building the new consensus sequence, a minimum mapping quality of 1 and a minimum neighboring quality of 1 were tolerated. These parameters were chosen to minimize the number of Ns in the new consensus. Given our long read length (76 bp), we did not expect false matches to be a problem. All sequences generated by either direct sequencing or assembling fragments from Illumina sequencing were submitted to GenBank under accession nos. GU378397 - GU378437 (*atpA*), GU378438 - GU378496 (*gyrB*), GU378497 - GU378524 (*pyrH*), GU378525 - GU378574 (*recA*) and GU378575 - GU378610 (*topA*).

Previously Determined *Vibrio* Gene Sequences

427 bp of the published *hsp60* sequences from fract-2006 (Hunt *et al.*, 2008a), invert-2007 (Chapter 2, this thesis) and particle-2007 (Chapter 3, this thesis) were used in population comparisons. Sequences for all named *Vibrio* species were obtained from GenBank using accession numbers provided by previously published phylogenetic analyses of *Vibrionaceae* (Le Roux *et al.*, 2004; Sawabe *et al.*, 2007; Thompson *et al.*, 2005a). Additionally, sequences with no close relative were blasted against the database (Oct, 22, 2009). Best blast hits to named species were added to analysis according to Table 5. These additional sequences were primarily from *V. breoganii*, which was characterized after publication of phylogenetic analyses used here.

Sequence Alignment and Phylogenetic Inference

Phylogenetic relationships for all loci used in this analysis were determined by maximum likelihood using PhyML v.2.4 (Guindon & Gascuel, 2003b). The GTR substitution model and four rate categories (parameters were estimated from the data) were used for the phylogenetic analysis. The iTOL tool (Letunic & Bork, 2007) was used to visualize the distribution of isolates from each study in Fig. 1.

Environmental Correlation Analysis

A table of total number of isolates from each population was created based on the number of isolates obtained from each environmental category across studies. These were transformed into percentage of total in each environmental category. Cluster (Eisen *et al.*, 1998) was used to normalize by populations and environmental categories. Complete linkage hierarchical clustering on populations (“genes”) was done using uncentered correlations. The relative proportion of each population across environmental categories was visualized cluster using TreeView (Saldanha, 2004).

REFERENCES

- Broughton, L.C. & Gross, K.L. (2000) Patterns of diversity in plant and soil microbial communities along a productivity gradient in a Michigan old-field. *Oecologia*, **125**, 420-427.
- Eckburg, P.B., Bik, E.M., Bernstein, C.N., Purdom, E., Dethlefsen, L., Sargent, M., Gill, S.R., Nelson, K.E., & Relman, D.A. (2005) Diversity of the human intestinal microbial flora. *Science*, **308**, 1635-1638.

- Eisen, M.B., Spellman, P.T., Brown, P.O., & Botstein, D. (1998) Cluster analysis and display of genome-wide expression patterns. *Proceedings of the National Academy of Sciences of the United States of America*, **95**, 14863-14868.
- Fraser, C., Alm, E.J., Polz, M.F., Spratt, B.G., & Hanage, W.P. (2009) The Bacterial Species Challenge: Making Sense of Genetic and Ecological Diversity. *Science*, **323**, 741-746.
- Furlong, M.A., Singleton, D.R., Coleman, D.C., & Whitman, W.B. (2002) Molecular and culture-based analyses of prokaryotic communities from an agricultural soil and the burrows and casts of the earthworm *Lumbricus rubellus*. *Applied and Environmental Microbiology*, **68**, 1265-1279.
- Gevers, D., Cohan, F.M., Lawrence, J.G., Spratt, B.G., Coenye, T., Feil, E.J., Stackebrandt, E., Van de Peer, Y., Vandamme, P., Thompson, F.L., & Swings, J. (2005) Re-evaluating prokaryotic species. *Nature Reviews Microbiology*, **3**, 733-739.
- Guindon, S. & Gascuel, O. (2003) A simple, fast, and accurate algorithm to estimate large phylogenies by maximum likelihood. *Systematic Biology*, **52**, 696-704.
- Horner-Devine, M.C. & Bohannan, B.J.M. (2006) Phylogenetic clustering and overdispersion in bacterial communities. *Ecology*, **87**, S100-S108.
- Huber, J.A., Mark Welch, D., Morrison, H.G., Huse, S.M., Neal, P.R., Butterfield, D.A., & Sogin, M.L. (2007) Microbial population structures in the deep marine biosphere. *Science*, **318**, 97-100.
- Hubert, C., Loy, A., Nickel, M., Arnosti, C., Baranyi, C., Bruchert, V., Ferdelman, T., Finster, K., Christensen, F.M., de Rezende, J.R., Vandieken, V., & Jorgensen, B.B. (2009) A Constant Flux of Diverse Thermophilic Bacteria into the Cold Arctic Seabed. *Science*, **325**, 1541-1544.
- Hunt, D.E., David, L.A., Gevers, D., Preheim, S.P., Alm, E.J., & Polz, M.F. (2008a) Resource partitioning and sympatric differentiation among closely related bacterioplankton. *Science*, **320**, 1081-1085.
- Hunt, D.E., Gevers, D., Vahora, N.M., & Polz, M.F. (2008b) Conservation of the chitin utilization pathway in the Vibrionaceae. *Applied and Environmental Microbiology*, **74**, 44-51.
- Johnson, Z.I., Zinser, E.R., Coe, A., McNulty, N.P., Woodward, E.M.S., & Chisholm, S.W. (2006) Niche partitioning among *Prochlorococcus* ecotypes along ocean-scale environmental gradients. *Science*, **311**, 1737-1740.
- Kaller, M., Lundeberg, J., & Ahmadian, A. (2007) Arrayed identification of DNA signatures. *Expert Review of Molecular Diagnostics*, **7**, 65-76.
- Koeppel, A., Perry, E.B., Sikorski, J., Krizanc, D., Warner, A., Ward, D.M., Rooney, A.P., Brambilla, E., Connor, N., Ratcliff, R.M., Nevo, E., & Cohan, F.M. (2008) Identifying the fundamental units of bacterial diversity: A paradigm shift to incorporate ecology into bacterial systematics. *Proceedings of the National Academy of Sciences*, **105**, 2504-2509.
- Lane, D.J. (1991). 16S/23S rRNA sequencing. In *Nucleic Acid Techniques in Bacterial Systematics* (eds E. Stackebrandt & M. Goodfellow), pp. 115-175. Wiley & Sons, Chinchester.

- Le Roux, F., Gay, M., Lambert, C., Nicolas, J.L., Gouy, M., & Berthe, F. (2004) Phylogenetic study and identification of *Vibrio splendidus* -related strains based on *gyrB* gene sequences. *Diseases of Aquatic Organisms*, **58**, 143-150.
- Letunic, I. & Bork, P. (2007) Interactive Tree Of Life (iTOL): an online tool for phylogenetic tree display and annotation. *Bioinformatics*, **23**, 127-128.
- Ley, R.E., Peterson, D.A., & Gordon, J.I. (2006) Ecological and evolutionary forces shaping microbial diversity in the human intestine. *Cell*, **124**, 837-848.
- Li, M., Zhou, H.K., Hua, W.Y., Wang, B.H., Wang, S.Y., Zhao, G.P., Li, L.J., Zhao, L.P., & Pang, X.Y. (2009) Molecular diversity of *Bacteroides* spp. in human fecal microbiota as determined by group-specific 16S rRNA gene clone library analysis. *Systematic and Applied Microbiology*, **32**, 193-200.
- Lozupone, C. & Knight, R. (2005) UniFrac: a new phylogenetic method for comparing microbial communities. *Applied and Environmental Microbiology*, **71**, 8228-8235.
- Macian, M.C., Ludwig, W., Aznar, R., Grimont, P.A.D., Schleifer, K.H., Garay, E., & Pujalte, M.J. (2001) *Vibrio lentus* sp nov., isolated from Mediterranean oysters. *International Journal of Systematic and Evolutionary Microbiology*, **51**, 1449-1456.
- Martiny, J.B.H., Bohannan, B.J.M., Brown, J.H., Colwell, R.K., Fuhrman, J.A., Green, J.L., Horner-Devine, M.C., Kane, M., Krumins, J.A., Kuske, C.R., Morin, P.J., Naeem, S., Ovreas, L., Reysenbach, A.L., Smith, V.H., & Staley, J.T. (2006) Microbial biogeography: putting microorganisms on the map. *Nature Reviews Microbiology*, **4**, 102-112.
- Okada, K., Iida, T., Kita-Tsukamoto, K., & Honda, T. (2005) *Vibrios* commonly possess two chromosomes. *Journal of Bacteriology*, **187**, 752-757.
- Saldanha, A.J. (2004) Java Treeview-extensible visualization of microarray data. *Bioinformatics*, **20**, 3246-3248.
- Sawabe, T., Kita-Tsukamoto, K., & Thompson, F.L. (2007) Inferring the evolutionary history of *vibrios* by means of multilocus sequence analysis. *Journal of Bacteriology*, **189**, 7932-7936.
- Thompson, C.C., Thompson, F.L., Vicente, A.C.P., & Swings, J. (2007) Phylogenetic analysis of *vibrios* and related species by means of *atpA* gene sequences. *International Journal of Systematic and Evolutionary Microbiology*, **57**, 2480-2484.
- Thompson, F.L., Gevers, D., Thompson, C.C., Dawyndt, P., Naser, S., Hoste, B., Munn, C.B., & Swings, J. (2005a) Phylogeny and molecular identification of *vibrios* on the basis of multilocus sequence analysis. *Applied and Environmental Microbiology*, **71**, 5107-5115.
- Thompson, F.L., Iida, T., & Swings, J. (2004) Biodiversity of *vibrios*. *Microbiology and Molecular Biology Reviews*, **68**, 403-+.
- Thompson, F.L., Thompson, C.C., Li, Y., Gomez-Gil, B., Vandenberghe, J., Hoste, B., & Swings, J. (2003a) *Vibrio kanaloae* sp nov., *Vibrio pomeroyi* sp nov and *Vibrio chagasii* sp nov., from sea water and marine animals. *International Journal of Systematic and Evolutionary Microbiology*, **53**, 753-759.

- Thompson, F.L., Thompson, C.C., & Swings, J. (2003b) *Vibrio tasmaniensis* sp nov., isolated from Atlantic Salmon (*Salmo salar* L.). *Systematic and Applied Microbiology*, **26**, 65-69.
- Thompson, J.D., Gibson, T.J., Plewniak, F., Jeanmougin, F., & Higgins, D.G. (1997) The CLUSTAL_X windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. *Nucleic Acids Research*, **25**, 4876-4882.
- Thompson, J.R., Pacocha, S., Pharino, C., Klepac-Ceraj, V., Hunt, D.E., Benoit, J., Sarma-Rupavtarm, R., Distel, D.L., & Polz, M.F. (2005b) Genotypic diversity within a natural coastal bacterioplankton population. *Science*, **307**, 1311-1313.
- Venter, J.C., Remington, K., Heidelberg, J.F., Halpern, A.L., Rusch, D., Eisen, J.A., Wu, D.Y., Paulsen, I., Nelson, K.E., Nelson, W., Fouts, D.E., Levy, S., Knap, A.H., Lomas, M.W., Nealson, K., White, O., Peterson, J., Hoffman, J., Parsons, R., Baden-Tillson, H., Pfannkoch, C., Rogers, Y.H., & Smith, H.O. (2004) Environmental genome shotgun sequencing of the Sargasso Sea. *Science*, **304**, 66-74.
- Visick, K.L. & McFall-Ngai, M.J. (2000) An exclusive contract: Specificity in the *Vibrio fischeri* *Euprymna scolopes* partnership. *Journal of Bacteriology*, **182**, 1779-1787.
- Vos, M. & Didelot, X. (2009) A comparison of homologous recombination rates in bacteria and archaea. *Isme Journal*, **3**, 199-208.
- Webb, C.O., Ackerly, D.D., McPeck, M.A., & Donoghue, M.J. (2002) Phylogenies and community ecology. *Annual Review of Ecology and Systematics*, **33**, 475-505.

FIGURES AND TABLES:

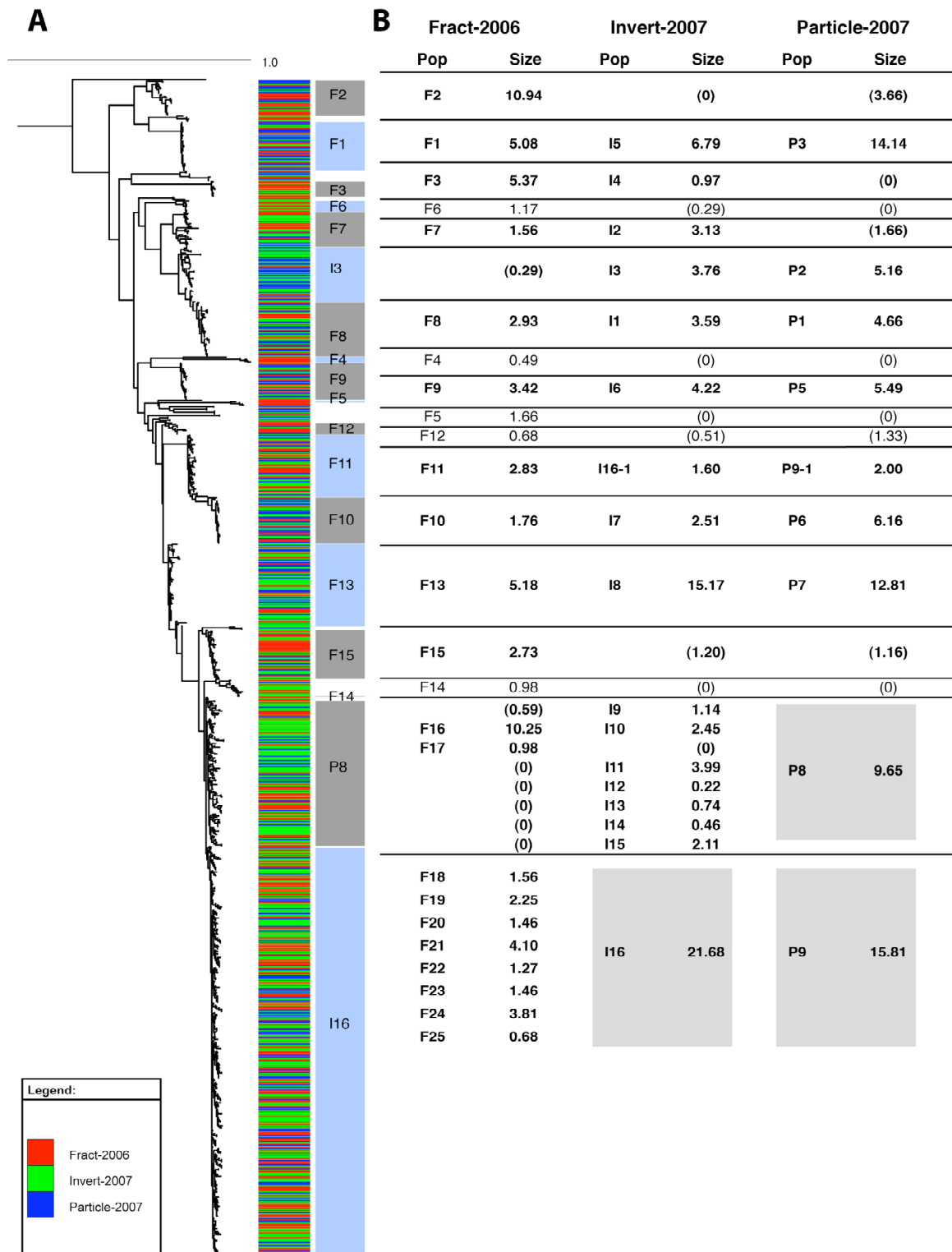


Fig. 1. (A) Phylogenetic relationship of populations predicted across three studies of *Vibrio* population structure (fract-2006, invert-2007 and particle-2007, explained in the text) inferred by maximum likelihood using *hsp60* sequences. The study each isolate was obtained from is identified by a red, blue or green color bar (left). Alternating gray and blue bars extend across the genetic breadth of the identified population, and are labeled according to a corresponding predicted population. **(B)** Summation of corresponding populations across all studies. Populations (represented by abbreviated study name and population numbers from each respective study) and the relative size of the population (as a percent of the total strains obtained in each study) are listed below the horizontal lines separating populations. Abbreviations as followed: Pop, population number with the single letter study designation in front; Size, population size as a percentage of total isolates in that study; F, fract-2006; I, invert-2007; P, particle-2007. Population sizes are shown in parenthesis when a confident ecological association could not be predicted. Large populations, which make at least 2% of total in one of the studies, are in bold. Sub-populations are indicated by multiple population numbers from the same study under the same horizontal line. Populations that span multiple sub-populations contained in a gray box.

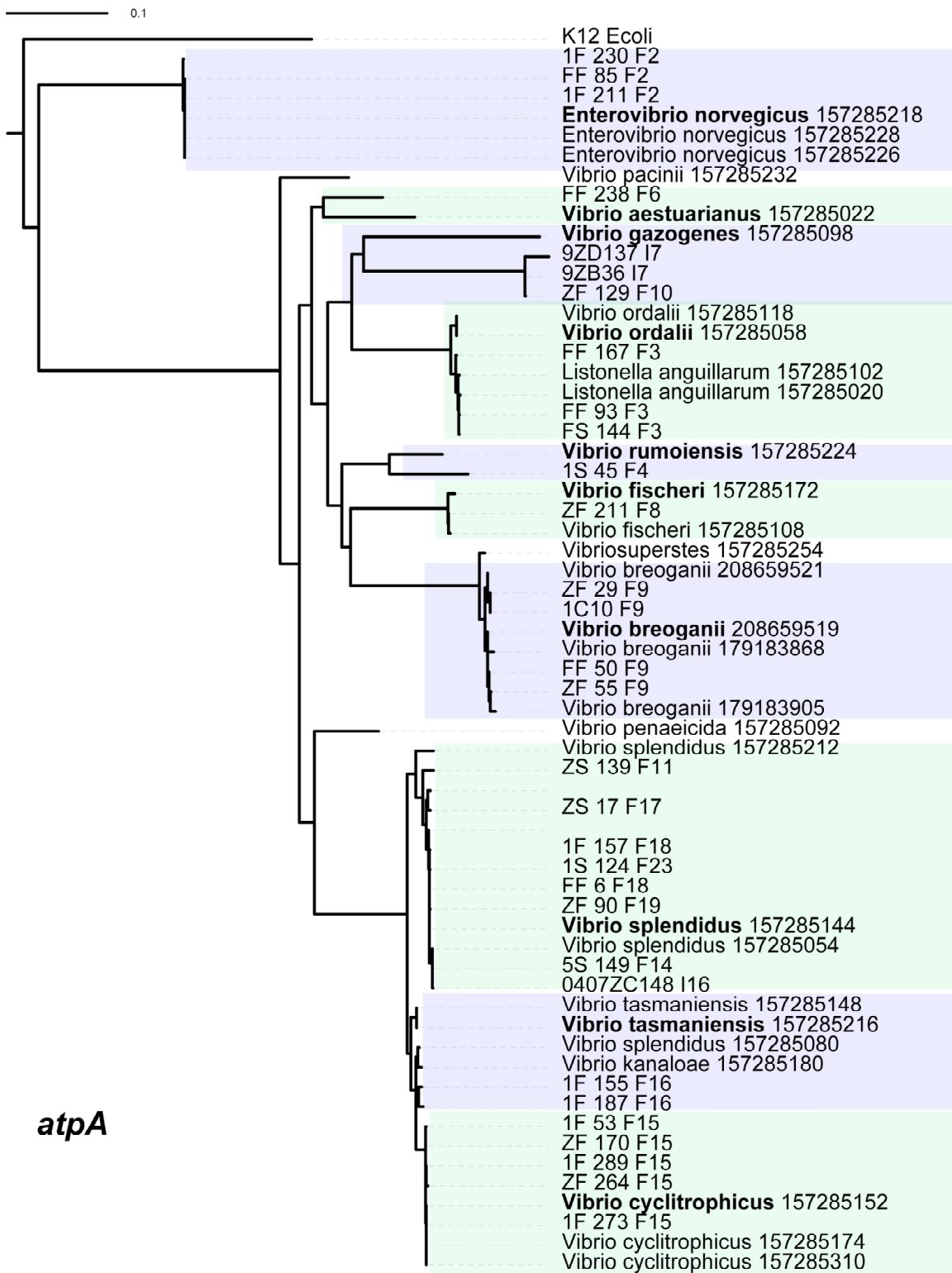


Fig. 2. Maximum likelihood estimation of phylogenetic relationship between named species and population representatives using partial *atpA* sequences. Clusters are alternatively shaded blue and green, and an associated type strains for each cluster is in bold. Species names followed by the NCBI GI number are provided for all sequences from NCBI. Population representatives are named using the isolate name followed by the study (using a single letter designation for the study) and population number.

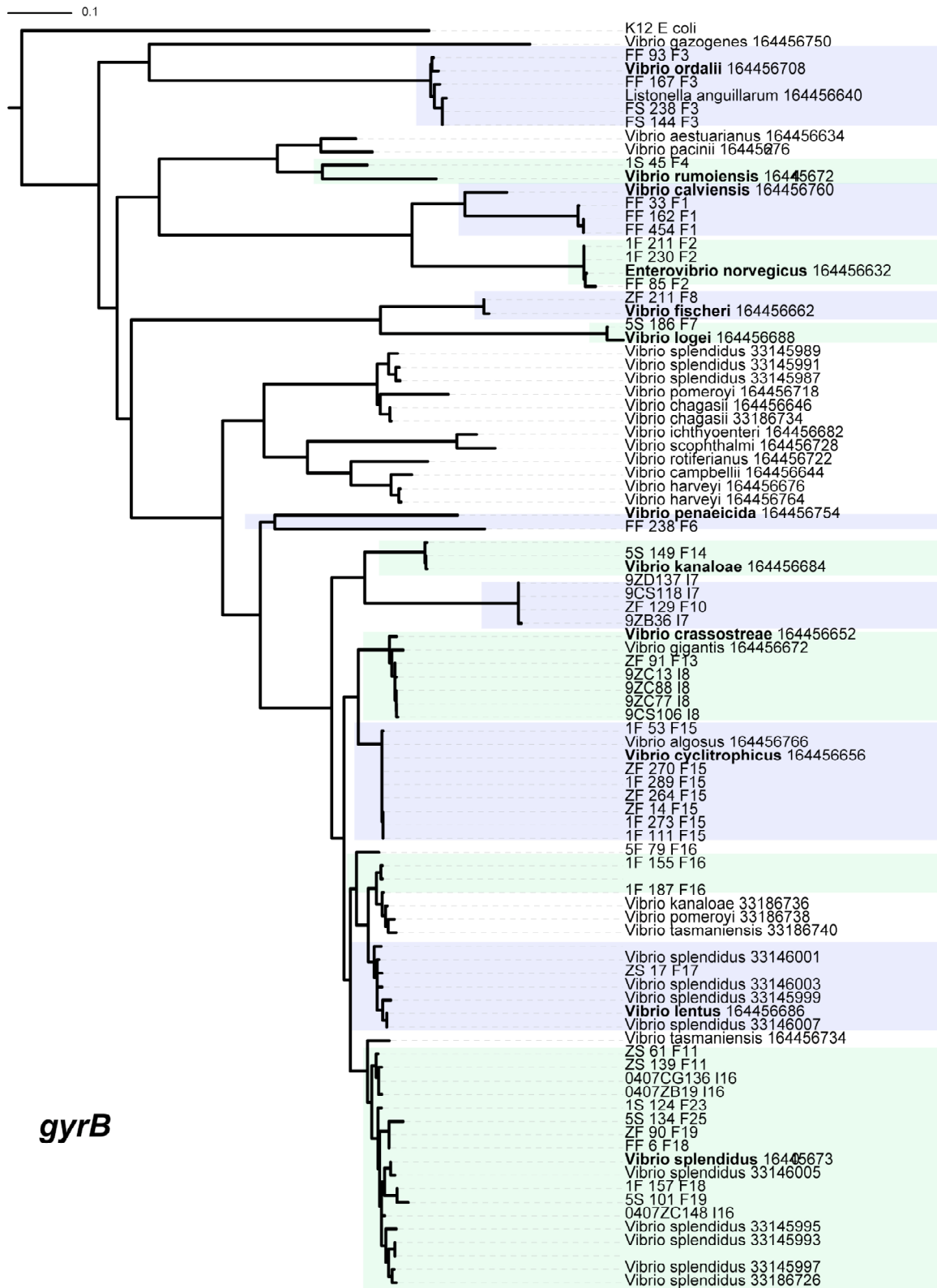


Fig. 3. Maximum likelihood estimation of phylogenetic relationship between named species and population representatives using partial *gyrB* sequence. Clusters are alternatively shaded blue and green, and an associated type strains for each cluster is in bold. Species names followed by the NCBI GI number are provided for all sequences from NCBI. Population representatives are named using the isolate name followed by the study (using a single letter designation for the study) and population number.

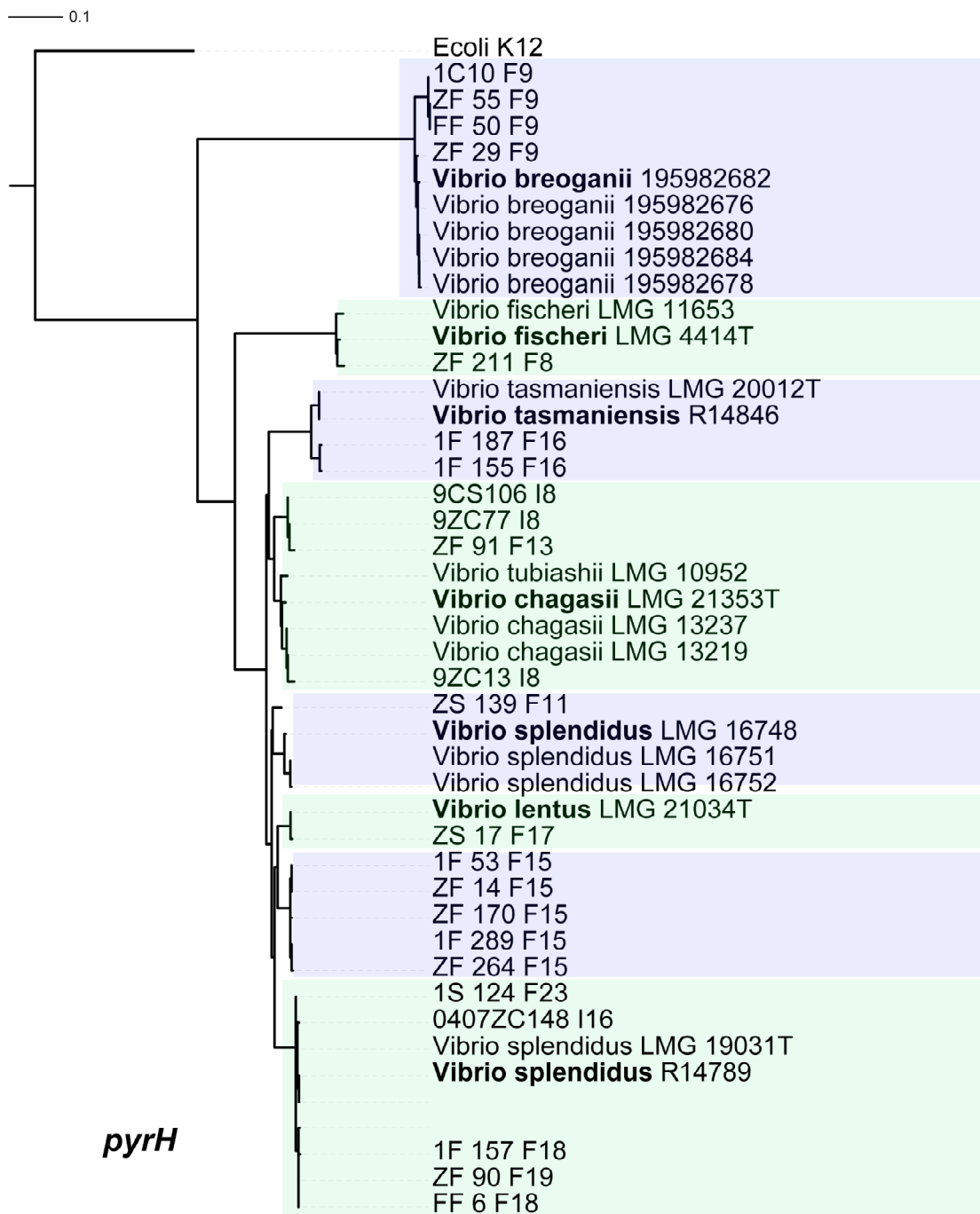


Fig. 4. Maximum likelihood estimation of phylogenetic relationship between named species and population representatives using partial *pyrH* sequence. Clusters are

alternatively shaded blue and green, and an associated type strains for each cluster is in bold. Species names followed by the NCBI GI number are provided for all sequences from NBCI. Population representatives are named using the isolate name followed by the study (using a single letter designation for the study) and population number.

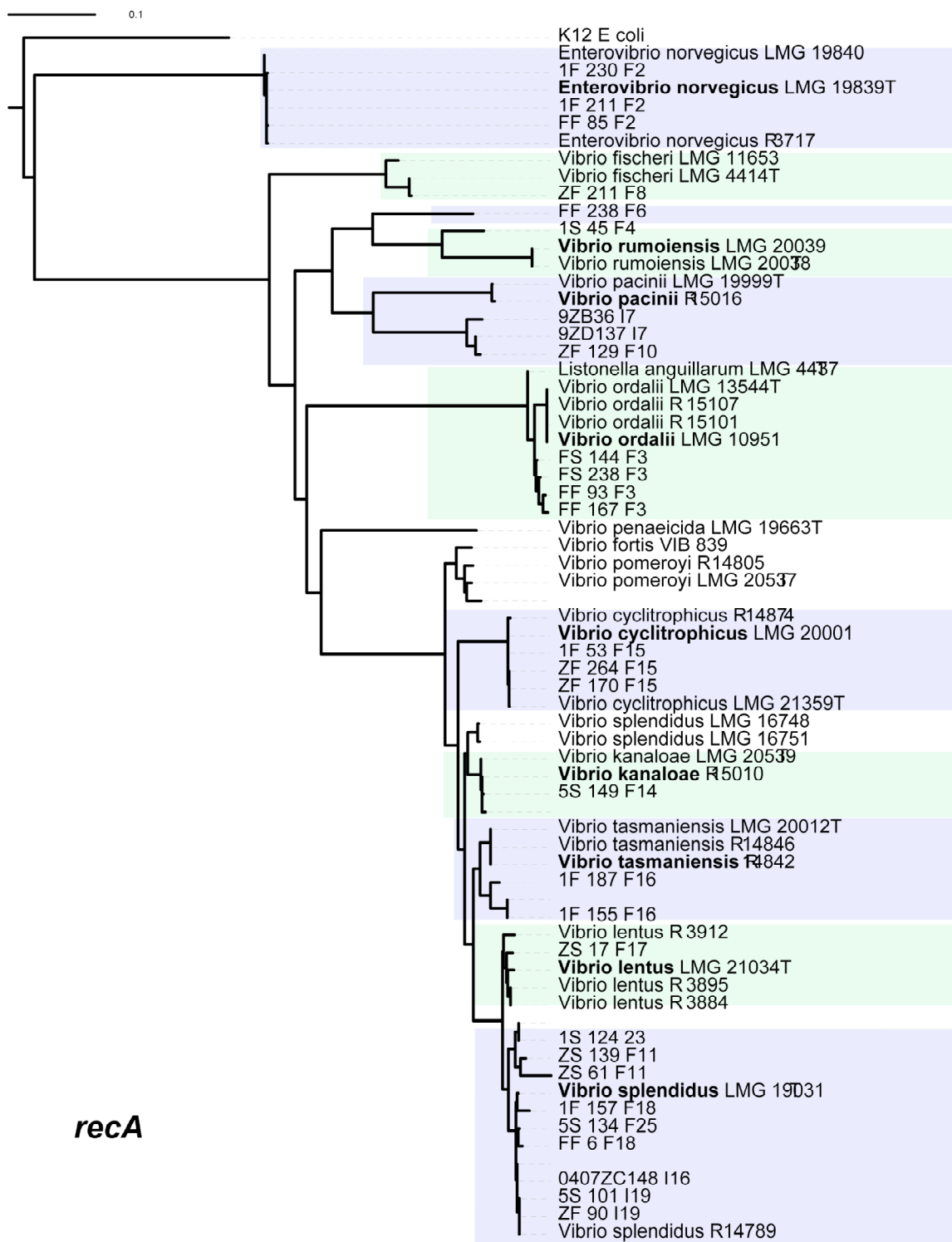


Fig. 5. Maximum likelihood estimation of phylogenetic relationship between named species and population representatives using partial *recA* sequences. Clusters are alternatively shaded blue and green, and an associated type strains for each cluster is in bold. Species names followed by the strain name are provided for all sequences from NBCI. Population representatives are named using the isolate name followed by the study (using a single letter designation for the study) and population number.

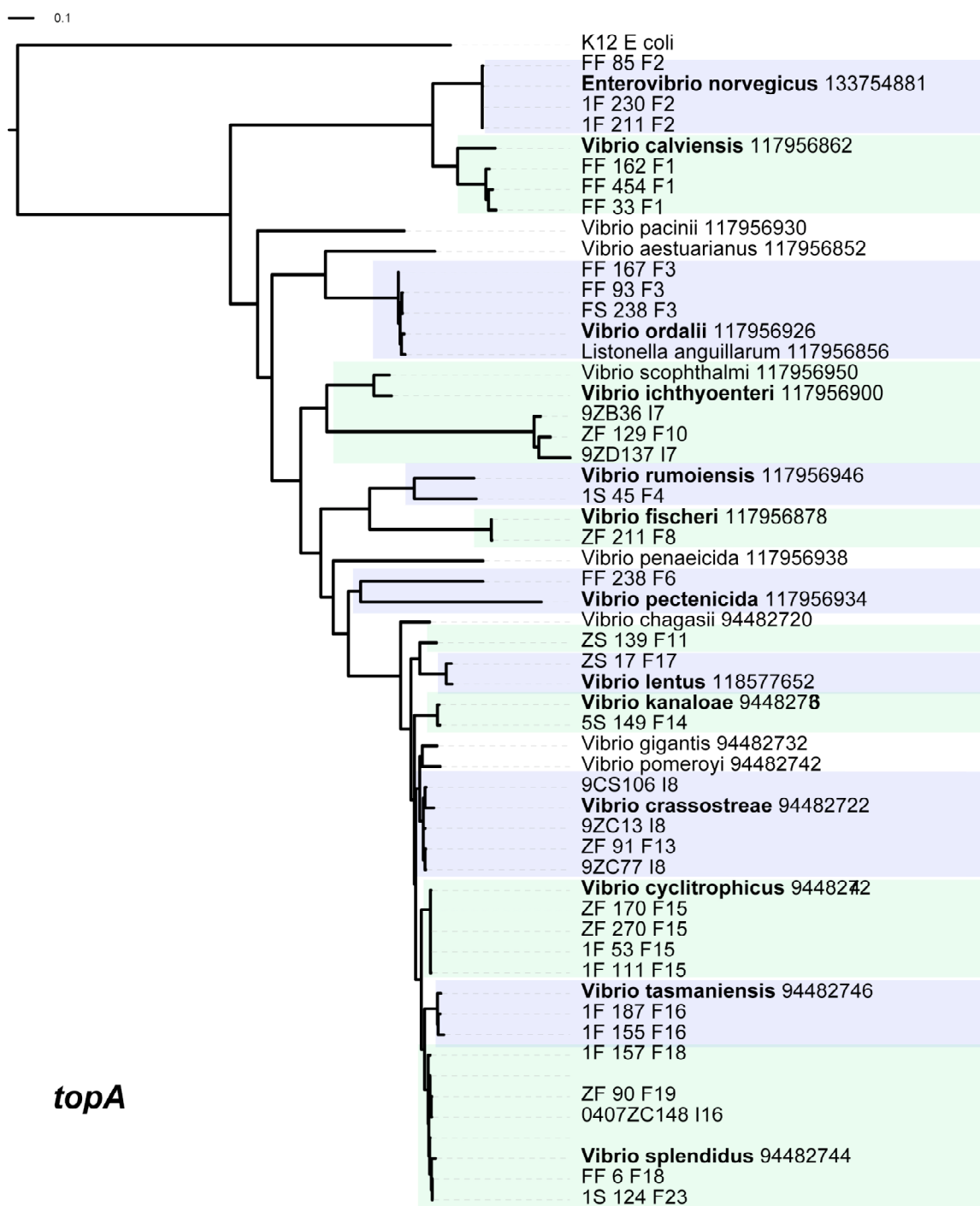


Fig. 6. Maximum likelihood estimation of phylogenetic relationship between named species and population representatives using partial *topA* sequences. Clusters are alternatively shaded blue and green, and an associated type strains for each cluster is in

bold. Species names followed by the NCBI GI number are provided for all sequences from NBCI. Population representatives are named using the isolate name followed by the study (using a single letter designation for the study) and population number.

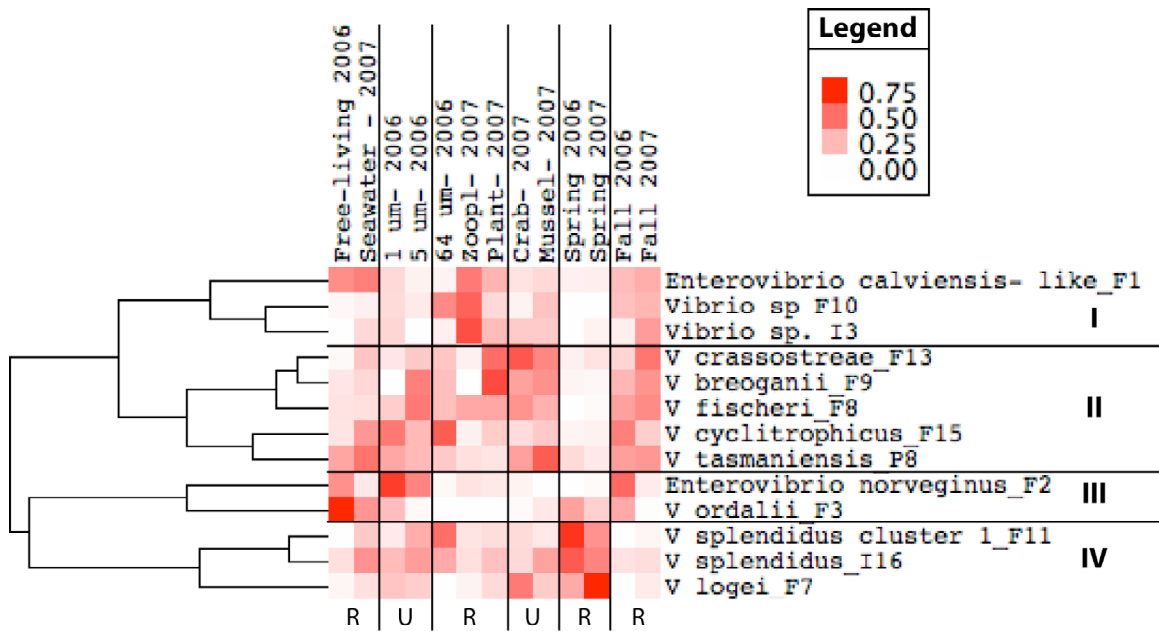


Fig. 7. Agglomerative hierarchical process used to generate a dendrogram (left) ordering populations by environmental similarity. Taxonomic groups are listed in a column on the right, consisting of the species names (if applicable), followed by population numbers from Table 2. Colored boxes represent the normalized distribution for each population across all habitats sampled in the three studies (values normalized by both population and environmental parameter to sum to an average value of 2.7416). Environmental parameters were grouped manually by similarity and are labeled with either an R (re-sampled) or U (unique), depending on whether they were re-sampled across years. Seasonal information is provided to the right of the dark vertical bar. Environmental profiles, or groups of populations with similar environmental distributions, are designated on the right by roman numerals I-IV. Environmental parameters are labeled as follows:

Free-living 2006, 1 μ M 2006 and 5 μ M 2006, corresponds to the free-living, 1 μ m and 5 μ M fractions from fract-2006, respectively; Zoopl 2006, the >64 μ M fraction from fract-2006; Zoopl 2007, living and dead zooplankton from particle-2007; Plant 2007, elongated and globular plant particles from particle-2007; Seawater 2007, whole seawater collected during invert-2007; Crab 2007 and Mussel 2007 are the summation of isolates from each invertebrate species from invert-2007; Fall 2007 and Spring 2007 is the sum of the invert-2007 and particle-2007 portions found in each season while Fall 2006 and Spring 2006 are the summation of isolates collected in each season from fract-2006.

Legend represents normalized values.

Table 1. Comparison of environmental conditions, parameters sampled and methodological information from three studies used in this meta-analysis.

Study/ Abbreviated Name	No. of Isolated Strains	Environmental Habitats (Total number)	Number of Replicate Samples	Spring Sampling Date; Temp	Fall Sampling Date; Temp	Genetic Loci	Number of Predicted Pops.
Hunt, et. al. 2008/ fract- 2006	1024	Sequential filtration through 64, 5, 1, 0.2 μ M filters (4)	4	04/28/06; 11 °C	09/06/06; 16 °C	hsp60	25
Chapter 2/ invert-2007	1753	Tissue and contents of Crabs and Mussels; pooled Zooplankton (10)	8	4/23/07- 05/01/07; 10 °C	9/25/07- 10/04/07; 14-16.5 °C	adk, hsp60, mdh	16
Chapter 3/ particle-2007	601	>64 μ m zooplankton and plant- derived particles (4)	8	4/30/07- 5/01/07; 7-10 °C	9/25/07- 9/28/07; 15-16.5 °C	adk, hsp60, mdh	9

Table 2. Species names and supporting phylogenetic information for modeled populations.

Population ¹	Size ²	Population Name	Genes Used ³		Notes
			Supported Clusters	Conflicting Clusters	
F2	L	<i>Enterovibrio norvegicus</i>	atpA, gyrB, recA, topA	None	Strong support for taxonomic name
F1	L	<i>Enterovibrio calviensis</i> - like	gyrB, topA,	None	Consistently close, but distinct
F3	L	<i>V. ordalii</i>	atpA, gyrB, recA, topA	None	Strong support for taxonomic name
F6	S	<i>Vibrio sp F6</i>	NA	atpA, gyrB, recA	Closest named relative varies by gene: <i>V. aesturanius</i> (atpA), <i>V. penaeicida</i> (gyrB), <i>V. rumoiensis</i> (recA)
F7	L	<i>Aliivibrio logei</i>	gyrB	None	Sequence information for only one gene; weak support
I3	L	<i>Aliivibrio sp. I3</i>	NA	NA	No sequence information
F8	L	<i>Aliivibrio fischeri</i>	atpA, gyrB, pyrH, recA, topA	None	Strong support for taxonomic name
F4	S	<i>V. rumoiensis</i> - like	atpA, gyrB, recA, topA	None	Consistently close, but distinct
F5	S	<i>Vibrio sp. F5</i>	NA	NA	No sequence information
F9	L	<i>V. breoganii</i>	atpA, pyrH	None	Strong support for taxonomic name
F13	L	<i>V. crassostreae</i>	gyrB, topA	PyrH with <i>V. chagasii</i> and <i>V. crassostreae</i> not close (95% by blast, but not on tree)	Despite conflict, this population seems close to <i>V. crassostreae</i>
F10	L	<i>Vibrio sp. F10</i>	NA	atpA, gyrB, recA	Closest named relative varies by gene: <i>V. pacinii</i> (recA), <i>V. kanaloae</i> (gyrB) and <i>V. gazogenes</i> (atpA)
F11	L	<i>V. splendidus</i> cluster 1	atpA, gyrB, pyrH, recA	topA	Close to <i>V. splendidus</i> , but both hsp60 and topA signals conflict
F12	S	<i>Vibrio sp. F12</i>	NA	NA	No sequence data available

F15	L	<i>V. cyclitrophicus</i>	atpA, gyrB, recA, topA	Some recombination	Strong support for taxonomic name
P8	L	<i>V. tasmaniensis</i>	atpA, gyrB, pyrH, recA, topA	NA	Strong support for taxonomic name
F17	S	<i>V. lentus</i>	gyrB, pyrH, recA, topA	hsp60 with <i>V.</i> <i>tasmaniensis</i>	Hunt et al (2008) show recombination at hsp60, otherwise loci support naming
I16	L	<i>V. splendidus</i>	atpA, gyrB pyrH, recA, topA	NA	Strong support for taxonomic name
F14	S	<i>V. kanaloae</i>	GyrB, recA	atpA	With <i>V. splendidus</i> at atpA, otherwise, good support taxonomic name.

¹ Populations correspond to groups of equivalent populations in Fig. 1.

² Small populations (S) are noted if population does not achieve greater than 2% total abundance in any study. Large populations (L) achieve greater than 2% total abundance and are used to assess reproducibility and specificity of associations.

³ Not all populations had sequence data available for taxonomically relevant loci for all genes, thus, this column provides information about which sequences were used to inform the species name and where there were conflicting signals with this naming.

Supported clusters show loci where the strains from the population clusters with named species, and conflicting clusters are shown where strains cluster with alternative named species.

Table 3. Percent of total isolates in each population associated with large (> 64 um) particles from both fract-2006 and particle-2007

Population	% fract-2006 ¹	% particle-2007 ²	Difference	Reproduced ³	Discordant
<i>V. breoganii</i>	4.64	5.49	0.85	x	
<i>Vibrio</i> sp F10	5.06	6.16	1.09	x	
<i>Aliivibrio fischeri</i>	3.38	4.66	1.28	x	
<i>V. crassostreae</i>	9.70	12.81	3.11	x	
<i>V. cyclitrophicus</i>	5.49	1.16	4.32	x	
<i>V. splendidus</i> cluster 2	27.85	22.30	5.55	x	
<i>V. splendidus</i> cluster 1	7.59	2.00	5.60	x	
<i>V. tasmaniensis</i>	2.95	9.65	6.70		x
<i>Aliivibrio</i> sp. 13	0.42	7.65	7.23		x
<i>Enterovibrio calviensis</i> - like	1.27	14.14	12.88		x

¹ Isolates found in each population associated with large particle fraction as a percentage of total isolates from large (> 64 um) particles, from fract-2006

² Isolates found in each population associated with large particle fraction as a percentage of total isolates from large (> 64 um) particles, from particle-2007

³ Populations are characterized as reproduced if the difference in abundance of isolates associated with large particles (as a percent of total isolates from the large particle fraction) between studies is less than 6%. Populations with abundance differences greater than 6% marked as discordant.

Table 4. List of sequences used in multi-locus gene sequence assembly using unassembled Illumina sequenced fragments.

Strain	Locus	NCBI GI No. or Accession No.
1F_211	atpA	157285228
1F_230	atpA	157285228
FF_85	atpA	157285228
FF_167	atpA	157285118
FF_93	atpA	157285118
FS_144	atpA	157285118
FS_238	atpA	157285118
1S_45	atpA	157285224
FF_238	atpA	157285022
5S_168	atpA	157285018
ZF_211	atpA	157285172
1C10	atpA	208659521
FF_50	atpA	208659521
ZF_29	atpA	208659521
ZF_55	atpA	208659521
9CS118	atpA	157285082
9ZB36	atpA	157285082
9ZD137	atpA	157285082
ZF_129	atpA	157285082
0407CG136	atpA	157285212

0407ZB19	atpA	157285212
0407ZC148	atpA	157285212
ZS_139	atpA	157285212
ZS_61	atpA	157285212
9CS106	atpA	157285174
9ZC13	atpA	157285174
9ZC40	atpA	157285174
9ZC77	atpA	157285174
ZF_91	atpA	157285174
5S_149	atpA	157285180
1F_155	atpA	179184105
1F_187	atpA	179184105
FF_500	atpA	179184105
ZS_17	atpA	179184105
1F_157	atpA	157285054
1S_206	atpA	157285054
FF_6	atpA	157285054
ZF_90	atpA	157285054
1S_124	atpA	157285054
1S_130	atpA	157285054
FF_238	gyrB	56409827
9CS118	gyrB	GU379148
9ZB36	gyrB	GU379148
9ZD137	gyrB	GU379148
ZF_129	gyrB	GU379148

0407CG136	gyrB	GU378456
0407ZB19	gyrB	GU378456
0407ZC148	gyrB	GU378456
ZS_139	gyrB	GU378456
9ZC13	gyrB	GU378470
9ZC40	gyrB	GU378470
9ZC77	gyrB	GU378470
5S_149	gyrB	82940537
1F_155	gyrB	82940539
1F_187	gyrB	82940539
5F_79	gyrB	82940539
FF_500	gyrB	82940539
ZS_17	gyrB	82940539
1S_206	gyrB	GU378459
1S_206	gyrB	GU378468
ZF_90	gyrB	33186728
1S_124	gyrB	33186728
1S_130	gyrB	33186728
0407CG136	pryH	195927663
0407ZB19	pryH	195927663
0407ZC148	pryH	195927663
ZS_139	pryH	195927663
ZS_61	pryH	195927663
5S_168	pyrH	156251952
ZF_211	pyrH	164608665

1C10	pyrH	195982676
FF_50	pyrH	195982676
ZF_29	pyrH	195982676
ZF_55	pyrH	195982676
9CS106	pyrH	195927655
9ZC13	pyrH	195927655
9ZC40	pyrH	195927655
9ZC77	pyrH	195927655
ZF_91	pyrH	195927655
1F_155	pyrH	195927657
1F_187	pyrH	195927657
FF_500	pyrH	195927657
ZS_17	pyrH	195927657
1F_157	pyrH	195927663
1S_206	pyrH	195927663
FF_6	pyrH	195927663
ZF_90	pyrH	195927663
1S_124	pyrH	195927663
1S_130	pyrH	195927663
FF_162	recA	223635952
FF_33	recA	223635952
FF_454	recA	223635952
1F_211	recA	58415330
1F_230	recA	58415330
FF_85	recA	58415330

FF_238	recA	157100332
5S_168	recA	GU380042
ZF_211	recA	GU380214
9CS118	recA	GU380087
9ZB36	recA	GU380087
9ZD137	recA	GU380087
ZF_129	recA	GU380087
0407CG136	recA	GU378551
0407ZB19	recA	GU378551
0407ZC148	recA	GU378551
ZS_61	recA	GU378551
9ZC13	recA	GU378553
9ZC40	recA	GU378553
ZF_91	recA	GU378553
1F_155	recA	58415533
1F_187	recA	58415533
FF_500	recA	58415533
ZS_17	recA	58415533
1S_206	recA	58415653
FF_6	recA	58415653
ZF_90	recA	58415653
1S_124	recA	58415653
1S_130	recA	58415653
FF_162	topA	117956862
FF_33	topA	117956862

FF_454	topA	117956862
1F_211	topA	133754885
1F_230	topA	133754885
FF_85	topA	133754885
FF_167	topA	117956926
FF_93	topA	117956926
FS_144	topA	117956926
FS_238	topA	117956926
1S_45	topA	117956946
FF_238	topA	117956852
5S_168	topA	117956902
ZF_211	topA	117956878
9CS118	topA	94482744
9ZB36	topA	94482744
9ZD137	topA	94482744
ZF_129	topA	94482744
0407CG136	topA	94482744
0407ZB19	topA	94482744
0407ZC148	topA	94482744
ZS_139	topA	94482744
ZS_61	topA	94482744
9CS106	topA	94482728
9ZC13	topA	94482728
9ZC40	topA	94482728
9ZC77	topA	94482728

ZF_91	topA	94482728
5S_149	topA	94482736
1F_155	topA	118577652
1F_187	topA	118577652
FF_500	topA	118577652
ZS_17	topA	118577652
1F_157	topA	94482744
1S_206	topA	94482744
FF_6	topA	94482744
ZF_90	topA	94482744
1S_124	topA	94482744
1S_130	topA	94482744

Table 5. NCBI GI numbers for *V. breoganii* sequences used in this analysis, which were not included in previously published *Vibrio* taxonomic analysis.

NCBI GI No.	Strain Name	Gene
179183868	CECT 7222	atpA
179183905	LMG 24484	atpA
179183937	C 5.5	atpA
179184888	CECT 7222	recA
179185171	C 5.5	recA
179185203	LMG 24484	recA
195982676	RD 15.11	pyrH
195982678	CMJ 13.7	pyrH
195982680	VB 16.3	pyrH
195982682	C 5.5	pyrH
195982684	RD2 G5	pyrH
208659519	C 4.15	atpA
208659521	CMJ 13.7	atpA
208659523	RD2B2	atpA
208659525	VB 16.3	atpA
208659527	C 4.15	recA
208659529	CMJ 13.7	recA
208659531	RD2B2	recA
208659533	VB 16.3	recA

Chapter Five

Conclusions and Future Work

Bacteria dominate the biomass of the ocean and play a crucial part in maintaining a healthy, functioning ecosystem. The ocean also harbors an extremely diverse community of different bacterial types (Giovannoni & Stingl, 2005; Huber *et al.*, 2007). In order to understand how this vast amount of genetic diversity is generated and maintained, fundamental functional units of bacterial diversity must be identified using careful observation and verification. Further work can be done to elucidate how these units self-organize under different environmental conditions. This thesis, by building on previous work, uses careful observation of the distribution of one bacterial family across a range of environmental categories to identify these units (populations) and probe further into principles that likely shape their self-assembly (e.g. growth on different substrates, relative growth on particles *in vitro*).

This thesis investigates whether populations of the gamma-proteobacterial family, *Vibrionaceae*, display highly specialized distribution across a wide range of environmental categories, such as marine invertebrates or particles in the water column. Additionally, it seeks to determine whether *in situ* population distributions can be attributable to a competitive advantage over other vibrio populations, through *in vitro* competition assays on native particles. Finally, populations from this thesis are compared to a previous analysis of vibrio population structure and named *Vibrio* species to determine whether populations are consistently predicted from associations with different environmental categories or correspond to named species. These experiments highlight processes that shape and maintain diversity in natural bacterial populations.

The members of the *Vibrio* genus were used as a model system because they are

an important part of the bacterioplankton community and extensive diversity has been documented in this group (Thompson *et al.*, 2005), making them amenable for use in my study. Many members are pathogenic to both humans and marine biota, and three significant human pathogens in this genus have been identified (Thompson *et al.*, 2004). *Vibrio* species are highly abundant in many coastal environments and are important in nutrient cycling, especially as degraders of the highly abundant marine polymer, chitin. Members of the *Vibrio* genus have been found at high densities in diverse environments within the water column of the coastal ocean. Studies have determined that vibrios can be found in microaggregates (Simon *et al.*, 2002), biofilms (Kwon *et al.*, 2002), sediments (Kopke *et al.*, 2005), associated with marine animals and macroalgae (Sawabe *et al.*, 2004; Thomson *et al.*, 2005) and within the water column. Vibrios are amenable to multilocus sequencing because they can be isolated on selective media from the environment and grown in culture. Thus, the *Vibrio* genus is a tractable model system to use in this study.

This thesis has three main components: 1.) observation of the environmental distribution of vibrio populations 2.) testing principles that influence the environmental distribution through experimentation and 3.) assessing the reproducibility of population associations over time and replicate samples. Incorporating all of these aspects into this thesis highlights the strengths of this approach.

This thesis was motivated by previous work investigating important aspects of bacterial population structure in the ocean and identifying functional units of bacterial diversity. A study of the genetic diversity of vibrios at the same sampling site used in this

thesis established a model of bacterial adaptation which identifies functional units of bacterial diversity (populations) using the genetic and ecological relationship of strains to form the basis of the unit (Hunt *et al.*, 2008). Using this method, Hunt, David *et al* (2008) identified twenty-five co-existing vibrio populations that have one of six different lifestyles (habitats). This study established that vibrios strongly partition resources in the water column, and highly specialized particle preferences may have a strong influence on their evolution.

Building upon this work, this thesis extended the analysis of vibrio populations, observing their distribution across marine invertebrates (Chapter 2). Because vibrios are known as commensals of marine invertebrates, the distribution of vibrio populations within and across these organisms likely has a strong impact on their evolution. Three different marine invertebrates, with different lifestyles were chosen; mussels, crabs and zooplankton. Both the crab and zooplankton are covered with a chitinous carapace that extends across all tissues exposed to the environment, including most of their gut and their gill tissue. In contrast, marine mussel tissue is not protected by a chitinous cover, which might provide a different surface for bacterial attachment. Differences in feeding strategy between the large invertebrates is also substantial, as mussels are sessile and filter feed particles from the seawater and crabs used in this study are omnivorous, but prefer to eat mussels. Vibrios were also isolated from tissue and tissue content samples of different areas of the respiratory and digestive tract of the large invertebrates, which are areas where the differences in environmental conditions might promote bacterial specialization.

In spite of the difference between the sampled environments, the environmental distribution of vibrios across marine invertebrates demonstrated that *Vibrio* populations show little specialization for specific host types or host tissue. The preponderance of generalist populations suggests that migration into invertebrate hosts is high and corresponding selective pressure may be low. Population structure between host individuals likely created by clonal expansion was significant in crab and zooplankton samples, while little structure was observed between mussel individuals. This suggests that metapopulation structure, where the population is subdivided into patches and individuals can disperse between patches, is important for bacteria colonizing animals.

Since previous studies showed a high degree of specialization for particles of different sizes in the water column, the association of vibrio populations with particles of different types was also investigated (Chapter 3). Particles of either zooplankton origin (living and dead) or plant-derived origin (long or globular) were picked by eye under a dissecting scope to isolate vibrios from very specific types of particles in the water column. Of the nine populations predicted from strains isolated from these four different particle types, seven populations showed a high degree of specialization. Two of the populations showed a very even distribution across particles of all types in this study, but with different seasonal distributions. Other populations showed a strong association with either zooplankton or plant-derived particles. This suggests that vibrio populations might have very specialized abilities to attach to and/or grow on the specific substrates found on each particle type.

Using environmental observations, this thesis extends our understanding of what

mechanisms determine vibrio population assembly onto particles in the environment through laboratory experimentation (Chapter 3). Three populations with environmental distributions that suggested the strongest degree of specialization for particles of each type were used to further investigate principles that determine their self-assembly onto particles of different types. Relative fitness between specialized vibrio populations was tested on native particles *in vitro* to determine the effects of competition in generating population structure. Extensive resource partitioning of both particle types and particle specific substrates was demonstrated, suggesting competitive exclusion as an important factor shaping distributions. Yet, competition experiments of specialized populations on native particles was not sufficient explain *in situ* distributions. Instead, either migration or the influence of other organisms (e.g. competitors, predators, viruses) is needed to fully explain the distribution of vibrio populations on particles of different types in the environment. All of this work highlights the important effects of migration and competition on the diversity of co-occurring vibrio populations in the coastal ocean.

Additionally, the reproducibility and specificity of *Vibrio* population associations was investigated across years by comparing a previous study of populations associated with water column particles to the associations determined from Chapters 2 and 3 of this thesis (Chapter 4). Reproducibility across years was high, but not all associations were reproduced. Specificity, as defined by consistently limited distribution across all environmental categories, was low, since many populations were found in almost all of the sampled categories. Finally, vibrio populations predicted from environmental associations were largely consistent with named *Vibrio* species. This information can be

used to refine our current thinking of the species concept in microbiology.

Although this thesis investigated the distribution of vibrio populations across a wide range of categories and used these observations to test reproducibility and the specificity of these populations, many questions remain. Future work should be done to address questions that could not be answered in this thesis. More observation of the association of vibrio populations with environmental categories not sampled in this thesis will provide more insight into areas of the environment that strongly influence vibrio diversity. Additional observations may provide insight into where or when vibrio populations that were not commonly observed in this study may become more abundant (environmental reservoir). If possible, extending the analysis to include different bacterial or eukaryotic microbial population might identify important relationships between vibrios and other non-vibrio microbial populations.

More work should be done to extend the results from competition assays. Additional experiments *in vitro* on native particles and experimentally defined particles such as brown algae substrates will help determine if a more defined niche can be identified for some populations. Competition assays in this thesis only provide information about changes in population distribution relative to other populations. However, it would be much more informative to know how the absolute abundance of vibrio populations change over time on samples of each type. Having an absolute number, rather than a relative value, would make it easier to identify synergistic or antagonistic effects. Additionally, adding both plant-particle specialists populations together in *in vitro* competition assays, along with the zooplankton specialist population,

may increase the relative advantage of plant-particle specialist populations over the zooplankton specialists if cross-feeding is indeed occurring between populations. Also, since the zooplankton specialist population used for *in vitro* studies seems to be relatively more abundant on living zooplankton, another competition assay, either using living zooplankton or dead zooplankton and one of the populations that are relatively more abundant on dead zooplankton may be more appropriate.

The relationship of the particle-specialist populations deserves more attention. Since it seems like *V. breoganii* requires *V. crassostreae* to achieve relatively high abundances on plant-derived particles, this suggests a mechanism of cross feeding of *V. breoganii* by *V. crassostreae*. Identifying difference in intermediate metabolites or expression of important genes between samples of each population grown alone and together on plant-derived particles might provide insight into the nature of their interaction. *Vibrio sp. F10* also seems to have an interesting relationship with living zooplankton as compared to other vibrio populations. Thus, the specifics of the interaction of this novel vibrio species with zooplankton could be investigated through careful laboratory work using cultured copepods or using fluorescence in situ hybridization (FISH) on zooplankton from the environment.

Assessing the reproducibility of vibrio population associations with environmental categories remains one of the largest unknowns. Replicating previous field studies as closely as possible will show how reproducibly populations and sub-populations partition into their respective habitats, especially if environmental conditions are as similar as possible to previously sampled conditions. Ideally, these studies could be

repeated many years in a row to identify changes in population structure, which, with careful observation, might be correlated with difference in environmental conditions. Information about how specialist populations, such as *V. ordalii* and *Vibrio* sp. F10, are adapted to their unique environment might be obtained by sequencing and analyzing the genome of these populations, especially in relation to other, more generalist populations. All of these additional experiments build on this work and will enhance our knowledge of the mechanisms creating and maintaining bacterial diversity in the environment.

REFERENCES

- Giovannoni, S.J. & Stingl, U. (2005) Molecular diversity and ecology of microbial plankton. *Nature*, **437**, 343-8.
- Huber, J.A., Mark Welch, D., Morrison, H.G., Huse, S.M., Neal, P.R., Butterfield, D.A., & Sogin, M.L. (2007) Microbial population structures in the deep marine biosphere. *Science*, **318**, 97-100.
- Hunt, D.E., David, L.A., Gevers, D., Preheim, S.P., Alm, E.J., & Polz, M.F. (2008) Resource partitioning and sympatric differentiation among closely related bacterioplankton. *Science*, **320**, 1081-1085.
- Kopke, B., Wilms, R., Engelen, B., Cypionka, H., & Sass, H. (2005) Microbial diversity in coastal subsurface sediments: a cultivation approach using various electron acceptors and substrate gradients. *Applied and Environmental Microbiology*, **71**, 7819-7830.
- Kwon, K.K., Lee, H.S., Jung, S.Y., Yim, J.H., Lee, J.H., & Lee, H.K. (2002) Isolation and identification of biofilm-forming marine bacteria on glass surfaces in Dae-Ho Dike, Korea. *Journal of Microbiology*, **40**, 260-266.
- Sawabe, T., Hayashi, K., Moriwaki, J., Thompson, F.L., Swings, J., Potin, P., Christen, R., & Ezura, Y. (2004) *Vibrio gallicus* sp nov., isolated from the gut of the French abalone *Haliotis tuberculata*. *International Journal of Systematic and Evolutionary Microbiology*, **54**, 843-846.
- Simon, M., Grossart, H.P., Schweitzer, B., & Ploug, H. (2002) Microbial ecology of organic aggregates in aquatic ecosystems. *Aquatic Microbial Ecology*, **28**, 175-211.
- Thompson, F.L., Iida, T., & Swings, J. (2004) Biodiversity of vibrios. *Microbiology and Molecular Biology Reviews*, **68**, 403-+.

- Thompson, J.R., Pacocha, S., Pharino, C., Klepac-Ceraj, V., Hunt, D.E., Benoit, J., Sarma-Rupavtarm, R., Distel, D.L., & Polz, M.F. (2005) Genotypic diversity within a natural coastal bacterioplankton population. *Science*, **307**, 1311-1313.
- Thomson, R., Macpherson, H.L., Riaz, A., & Birkbeck, T.H. (2005) *Vibrio splendidus* biotype 1 as a cause of mortalities in hatchery-reared larval turbot, *Scophthalmus maximus* (L.). *Journal of Applied Microbiology*, **99**, 243-250.